

Analisis Sentimen Terhadap Presidensi G20 2022 pada Media Sosial Twitter Menggunakan Metode Naïve Bayes

I Gusti Agung Indrawan¹, Dewa Ayu Indah Cahya Dewi^{2,*}, Ida Ayu Putu Ananda Wisdantini¹

¹Fakultas Teknologi dan Informatika, Prodi Teknik Informatika, Institut Bisnis dan Teknologi Indonesia, Denpasar, Indonesia

²urusan Teknik Elektro, Program Studi Sarjana Terapan Teknik Otomasi, Politeknik Negeri Bali, Badung, Indonesia

Email: ¹agung.indrawan@instiki.ac.id, ^{2,*}ayuindahcahyadewi@pnb.ac.id, ³dayuananda247@gmail.com

Email Penulis Korespondensi: ayuindahcahyadewi@pnb.ac.id

Abstrak—Twitter merupakan salah satu media sosial sebagai wadah yang cocok untuk mengemukakan pendapat dari masyarakat pada masa kini. Banyak masyarakat mengemukakan opininya melalui unggahan pada media sosial Twitter terhadap isu-isu yang menjadi trending topik. Salah satu isu yang sedang trending topik pada tahun 2022 adalah pelaksanaan kegiatan presidensi G20 yang diadakan di Bali, Indonesia. Isu tersebut, dapat memunculkan opini positif ataupun negatif dari masyarakat. Pada penelitian, akan dilakukan analisis sentimen terhadap pelaksanaan kegiatan presidensi G20 tahun 2022 dengan menggunakan Python Google Colab, RapidMiner Studio, dan Orange Data Mining. Pengumpulan data awal terdapat 24.840 data yang akan melalui tahap pre-processing text, pelabelan, pembagian, dan klasifikasi data menggunakan metode klasifikasi Naïve Bayes. Sehingga mendapatkan hasil klasifikasi sentimen positif 1.600 data (72,37%) dan sentimen negatif 611 data (27,63%). Berdasarkan hasil klasifikasi sentimen dapat disimpulkan bahwa masyarakat mendukung adanya kegiatan pelaksanaan kegiatan G20 tahun 2022 di Bali, Indonesia dilihat dari sentimen positif lebih banyak dibandingkan sentimen negatif. Klasifikasi Naïve Bayes memiliki performa yang cukup baik dalam mengklasifikasi topik yang diteliti, dimana didapatkan nilai akurasi sebesar 88.01%.

Kata Kunci: Analisis Sentimen; KTT G20 Indonesia; Twitter; Klasifikasi Naïve Bayes;

Abstract—Twitter is one of the social media as a suitable forum to express opinions from the community today. Many people express their opinions through posts on social media Twitter on issues that are trending topics. One of the trending topics in 2022 is the implementation of the G20 presidency held in Bali, Indonesia. This issue, can generate positive or negative opinions from the community. In the research, a sentiment analysis will be carried out on the implementation of G20 presidency activities in 2022 using Python Google Colab, RapidMiner Studio, and Orange Data Mining. In the initial data collection, there are 24,840 data that will go through the stages of text pre-processing, labeling, sharing, and data classification using the Naïve Bayes classification method. So that it obtained the results of the classification of positive sentiment 1,600 data (72.37%) and negative sentiment 611 data (27.63%). Based on the results of the sentiment classification, it can be concluded that the public supports the implementation of G20 activities in 2022 in Bali, Indonesia, seen from positive sentiment more than negative sentiment. The Naïve Bayes classification has a fairly good performance in classifying the topics studied, where an accuracy value of 88.01% is obtained.

Keywords: Sentiment Analysis; G20 Indonesia Summit; Twitter; Naïve Bayes Classifier; RapidMiner

1. PENDAHULUAN

G20 merupakan sebuah forum tingkat tinggi yang menghimpun kekuatan ekonomi negara maju dan negara berkembang yang menyatukan para pemimpin global untuk membahas isu-isu perekonomian dunia [1]. Kali ini Indonesia diberikan kesempatan untuk memegang sekaligus menjadi tuan rumah kegiatan G20. Kegiatan presidensi G20 ini akan dilaksanakan selama satu tahun, mulai dari 1 Desember 2021 hingga 30 November 2022. Tema dari presidensi G20 ini adalah “*Recover Together, Recover Stronger*”. Dimana dengan ini Indonesia mengajak seluruh dunia bersama-sama untuk memulihkan ekonomi setelah *Covid-19* [2].

Indonesia sebagai tuan rumah dilaksanakannya kegiatan G20 pada tahun 2022, tentu saja hal ini tidak luput dari pandangan atau opini masyarakat seluruh dunia terutama Indonesia. Masyarakat bebas memberikan opini terhadap isu apapun termasuk pada kegiatan pelaksanaan G20 tahun 2022. Banyaknya opini yang diberikan oleh masyarakat sehingga dapat memunculkan berbagai opini, baik itu opini positif ataupun negatif. Karena saat ini perkembangan media yang begitu pesat, dapat memunculkan banyak media *online* mulai dari media berita sampai dengan media sosial dan opini-opini tersebut bisa disampaikan dengan memberikan ulasan melalui situs media sosial yang merupakan salah satu wadah yang cocok untuk mengemukakan pendapat dari masyarakat pada masa kini. Contoh media sosial yang cocok sebagai wadah untuk menyampaikan opini seperti Twitter, Instagram, *Blog* pribadi, dan lain sebagainya.

Berdasarkan uraian diatas, maka dapat dilakukan sebuah pengolahan data dengan analisis sentimen. Setiap *tweet* yang diunggah oleh masyarakat pada aplikasi Twitter tidak mengandung makna yang selaras, karena hal itulah perlu dilakukannya analisis terhadap opini masyarakat di Twitter mengenai pelaksanaan G20 tahun 2022 dengan mengklasifikasi opini positif dan negatif. Sehingga hasil *output* dari dilakukannya analisis sentimen pada pelaksanaan kegiatan presidensi G20 tahun 2022 ini akan berdampak menjadikannya sebagai bahan evaluasi pada operasional untuk meningkatkan kualitas pengadaaan kegiatan atau *event* internasional di Indonesia selanjutnya.

Data yang akan diklasifikasi pada penelitian ini, diambil dengan *range* data dari bulan Agustus 2022 sampai dengan bulan Januari 2023. Dimana pada periode tersebut banyak masyarakat yang memberikan berbagai pendapat atau opininya mengenai Presidensi G20 tahun 2022, dilihat dari *trending* topik G20 yang dibahas oleh masyarakat Twitter. Penarikan data awal akan dilakukan pada media sosial Twitter. Dijelaskan dalam jurnal [3] pada tahun 2017 pengguna Twitter di seluruh dunia mencapai 328 juta, dibandingkan dengan periode yang sama di tahun sebelumnya meningkat sekitar 14%. Indonesia termasuk pengguna yang produktif di *Twitter*, disebutkan sebanyak 77% pengguna *Twitter* dan

menghasilkan mencapai 4,1 miliar *tweet* sepanjang 2016. Maka Indonesia dapat dikatakan salah satu pengguna *Twitter* terbesar di dunia. Metode yang akan digunakan pada penelitian ini yaitu *Naïve Bayes*. Karena metode ini mampu menghasilkan tingkat akurasi yang tinggi untuk mengklasifikasi ulasan pengguna kedalam opini positif dan negatif. *Naïve Bayes* juga merupakan klasifikasi yang paling sederhana [4].

Penelitian ini mengacu pada beberapa referensi penelitian terdahulu sebagai perbandingan. Penelitian pertama dilakukan oleh [5] yang berjudul “Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma *Naïve Bayes Classifier*”. Penelitian tersebut membahas tentang menganalisis pendapat vaksinasi COVID-19 di Indonesia yang dilakukan terhadap daya 3780 *tweet* yang berkaitan dengan vaksinasi menggunakan algoritma *Naïve Bayes Classifier*. Hasil dari penelitian tersebut, sebagian besar *tweet* memiliki sikap positif (60,3 %), sementara jumlah *tweet* yang netral (34,4 %) melebihi jumlah *tweet* yang menentang (5,4 %). Nilai akurasi yang dihasilkan sebesar 0,93 (93 %).

Penelitian kedua dilakukan oleh [6] yang berjudul “Analisis Sentimen Opini Publik Mengenai Covid-19 Pada Twitter Menggunakan Metode *Naïve Bayes* dan KNN”. Penelitian ini berfokus pada perbandingan hasil klasifikasi metode *Naïve Bayes* dan KNN, serta mengetahui kecenderungan opini masyarakat di Twitter. Penelitian ini menggunakan metode *Naïve Bayes*, memiliki nilai lebih tinggi dengan tingkat *accuracy* sebesar 63.21% dibandingkan dengan metode KNN yang memiliki tingkat *accuracy* sebesar 58.10% dan didapatkan 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 dari pada negatif yaitu 66.40% : 58.94%.

Penelitian ketiga dilakukan oleh [7] yang berjudul “Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi *Support Vector Machine (SVM)*”. Dimana dalam penelitiannya membahas tentang menganalisis sentimen opini pengguna layanan Indihome yang diunggah pada aplikasi Twitter dengan menggunakan metode *Support Vector Mecine (SVM)*. Penelitian ini menghasilkan *accuracy* 87%, *precision* 86%, *recall* 95%, *error rate* 13%, dan *f1-score* 90%. Dapat disimpulkan bahwa nilai positif sebesar 18,4 % dan hasil nilai negatif sebesar 81,6%.

Penelitian keempat dilakukan oleh [8] yang berjudul “Analisis Sentimen *Hate Speech* Pada Portal Berita Online Menggunakan *Support Vector Machine (SVM)*”. Penelitian ini membahas tentang pendeteksian kata-kata yang mengandung *hate speech* pada portal berita *online* dengan menggunakan pendekatan *Neural Language Processing*. Metode yang digunakan adalah *Support Vector Machine (SVM)*. Hasil analisis sentimen menunjukkan sebanyak 87 komentar untuk kategori *hate speech* (negatif) sementara untuk kategori tidak *hate speech* (positif) terdapat 105 komentar dengan rata-rata *accuracy* 53,88 % *recall* 49,69%, *precision* 48,77%, *Classification error* 46,12% dan *fmeasure* 49.23%.

Penelitian kelima dilakukan oleh [9] yang berjudul “Analisis Sentimen Opini Publik Terhadap Efek PSBB Pada Twitter Dengan Algoritma *Decision Tree*, KNN, dan *Naïve Bayes*”. Penelitian ini membahas tentang menganalisis pendapat pengguna jalan dan kendaraan yang kondisi di wilayahnya sedang diadakan PSBB serta membandingkan hasil klasifikasi dari algoritma *Decision Tree*, *Naïve Bayes*, dan KNN. Penelitian ini mendapatkan hasil *accuracy* dari *Decision Tree*, KNN, dan *Naïve Bayes* yaitu 83,3%, 80,80%, dan 80,03%. Hasil *precision* *Decision Tree*, K-NN, dan *Naïve Bayes* yaitu 81,06%, 82,72%, dan 87,54%. Sementara hasil *Recall* *Decision Tree*, K-NN, dan *Naïve Bayes* adalah 87,17%, 74,41%, dan 62,71%. Penelitian ini menyimpulkan bahwa *classifier Decision Tree* adalah *classifier* terbaik untuk digunakan dengan *dataset* media sosial karena memberikan prediksi yang lebih akurat dan tepat.

Dari kelima penelitian terdahulu diatas dapat disimpulkan bahwa penelitian-penelitian terdahulu memiliki perbedaan dari segi topik, metode, dan hasil penelitian. Perbedaan penelitian yang akan dilakukan dengan penelitian terdahulu adalah penelitian ini akan melakukan analisis sentimen pada pelaksanaan kegiatan presidensi G20 tahun 2022 dan data diambil dari media sosial Twitter, sehingga penelitian ini akan mengeluarkan *output* berupa total sentimen positif dan negatif dengan menggunakan metode *naïve bayes* sebagai klasifikasi data dan mengeluarkan hasil *wordcloud* sentiment positif dan negatif.

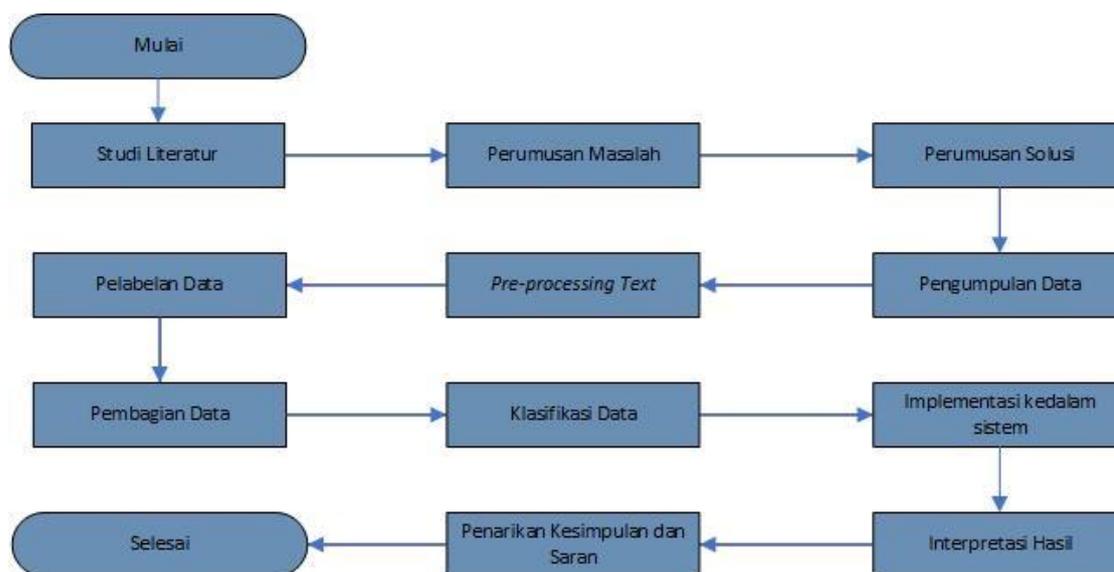
2. METODOLOGI PENELITIAN

2.1 Tempat dan Waktu Penelitian

Segala kegiatan penelitian ini dilakukan di kampus INSTIKI Indonesia yang terletak di Jl. Tukad Palerisan, No. 97, Panjer, Denpasar Selatan, Kota Denpasar, Bali. Dimana penelitian ini dilakukan pada semester ganjil tahun ajaran 2022/2023 mulai dari bulan Oktober sampai dengan semester genap tahun ajaran 2023/2024 bulan April.

2.2 Tahapan Penelitian

Tahapan penelitian ini berisi suatu jalannya atau alur penelitian yang akan dilakukan untuk mencapai tujuan penelitian. Berikut adalah alur tahapan penelitian yang dapat dilihat pada gambar 1.



Gambar 1. Alur Tahapan Penelitian

Berikut ini adalah penjelasan dari tahapan penelitian diatas, yaitu:

a. Studi Literatur

Studi literatur dilakukan untuk mendapatkan gambaran secara lengkap mengenai topik penelitian mulai dari pengumpulan data, menganalisis data, hingga mengimplementasi data dalam tools pada sistem. Untuk mendapatkan gambaran mengenai topik penelitian perlu didukung dengan teori-teori yang terkait dengan topik penelitian. Teori-teori terkait didapatkan dari jurnal, artikel, dan karya ilmiah yang ada pada internet, serta rujukan dari buku.

b. Perumusan Masalah

Perumusan masalah dilakukan oleh penulis untuk mengidentifikasi permasalahan awal, sehingga permasalahan tersebut dapat dijadikan sebagai bahan penelitian. Perumusan masalah pada penelitian ini adalah bagaimana menganalisis sentimen terhadap pelaksanaan kegiatan Presidensi G20 tahun 2022 pada aplikasi *Twitter* menggunakan metode *Naïve Bayes*.

c. Perumusan Tujuan

Tujuan dilakukannya penelitian adalah untuk menganalisis sentimen berbasis *text mining* terhadap pelaksanaan kegiatan Presidensi G20 tahun 2022 pada media sosial *Twitter* khususnya pada bulan Agustus 2022 hingga bulan Januari 2023.

d. Pengumpulan Data

Pengumpulan data dilakukan dengan cara meng-*crawling* data *Twitter* melalui *google colab* dengan menggunakan bahasa pemrograman *Python*. Pengumpulan data dilakukan dengan rentang data dari bulan Agustus 2022 hingga Januari 2023.

e. *Pre-processing Text*

Pre-processing text untuk membersihkan data-data yang sudah dikumpulkan sebelumnya agar data tersebut dapat dengan mudah dianalisis dan dapat memberikan data yang berkualitas.

f. Pelabelan Data

Pelabelan data dilakukan untuk mengetahui kelas-kelas pada suatu data. Dimana pelabelan digunakan untuk menentukan data tersebut masuk kedalam kelas positif ataupun negatif.

g. Pembagian Data

Pembagian data dilakukan untuk membagi data yang akan dianalisis menjadi data latih atau *training* dan data uji atau *testing*.

h. Klasifikasi Data

Klasifikasi data dilakukan untuk mengklasifikasi data yang sudah melewati beberapa tahap sebelumnya dengan menggunakan metode *naïve bayes* yang akan menghasilkan prediksi sentimen yang diuji dan tingkat *performance* dari klasifikasi data menggunakan *naïve bayes*.

i. Implementasi dalam Sistem

Implementasi dalam sistem akan dijelaskan pada bab 4 yang dimana pengolahan data akan dilakukan dengan memanfaatkan *tools-tools* yang ada pada *software RapidMiner Studio* dan *Orange Data Mining*.

j. Interpretasi Hasil

Interpretasi hasil merupakan sebuah penjelasan mengenai hasil yang didapatkan dari pengolahan data sentimen yang sudah dilakukan sebelumnya.

k. Penarikan Simpulan dan Saran

Kesimpulan akan berisi dari *point-point* penting dari hasil pengolahan data yang dilakukan pada penelitian ini, yang ditulis secara ringkas. Dan saran akan berisi saran-saran dalam pengolahan data dengan menggunakan metode dan media pengolahan data yang berbeda.

2.3 Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan untuk mempermudah dalam penyusunan laporan. Dalam penelitian ini, pengumpulan data menggunakan metode *crawling* data pada Twitter. Dimana data yang di *crawling* pada penelitian ini berasal dari Twitter, dikarenakan Twitter lebih unggul sebagai penyalur informasi atau berita tercepat dibandingkan dengan media sosial yang lainnya. Ketika terdapat topik baru yang muncul atau sedang *trending* dan menarik perhatian, maka akan ada banyak sekali opini masyarakat yang muncul di Twitter [10]. Penarikan data *tweet* hanya *tweet* yang berbahasa indonesia dan mengandung atau membahas tentang presidensi G20 tahun 2022 di Indonesia.



Gambar 2. Contoh *Tweet* Tentang Presidensi G20 Tahun 2022

3. HASIL DAN PEMBAHASAN

3.1 Proses Crawling Data

Penarikan data Twitter dilakukan melalui *Google Colab* menggunakan bahasa pemrograman *Python* yang memanfaatkan *library* “*snTwitter*”, kemudian mengisikan informasi berupa jumlah data yang ingin diambil, *keyword*, rentang bulan, tanggal dan tahun data *tweet* yang akan diambil.

Tabel 1. Hasil *Crawling* Data

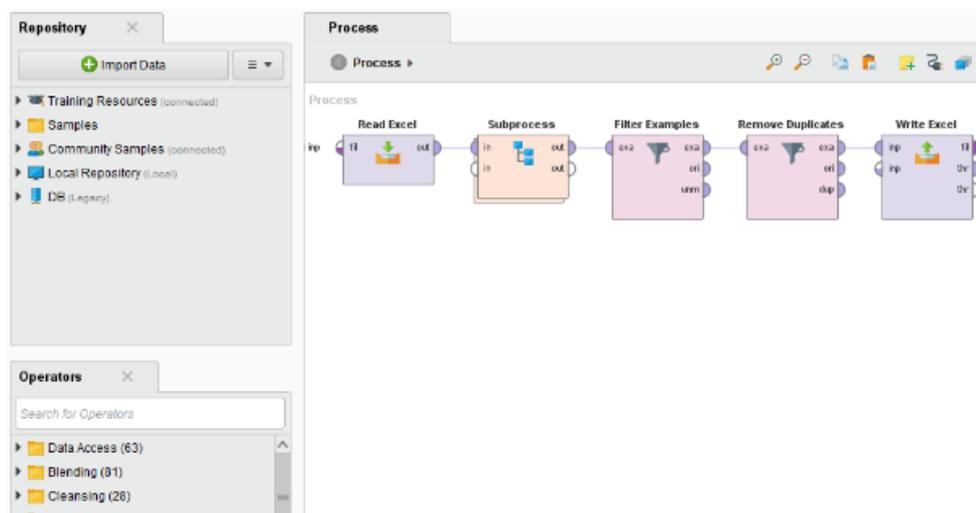
No	Text
1	Indonesia akan menjadi tuan rumah KTT G20 untuk pertama kalinya dan dijadwalkan berlangsung pada 15-16 November 2022 di Nusa Dua, Kabupaten Badung, Provinsi Bali. #g20Indonesia https://t.co/fNcGtRbReI
2	Pilar Presidensi G20 Indonesia 2022 dan Upaya Perdamaian Dunia https://t.co/OCW5pEtolz
3	Indonesia memegang Presidensi G20 Tahun 2022. Isu prioritas dalam G20 adalah upaya mengendalikan perubahan iklim yang berkontribusi kepada pemulihan yang berkelanjutan. https://t.co/IJZJCJildE #KilasBalikInsight
....	
24.840	Dengan KTT G20 Indonesia menerima Investasi besar untuk menyerap tenaga kerja https://t.co/NjkH32oLGM #KTTG20

Pada tabel 1 diatas merupakan data yang diperoleh pada proses *crawling* data dengan *keyword* “Pelaksanaan G20 Indonesia” dari tanggal 01-07 Agustus, dimana pada penelitian ini *crawling* data dilakukan per-minggu dari bulan Agustus 2022 hingga Januari 2023 dan di setiap minggu-nya menggunakan lima *keyword* untuk mengumpulkan data, diantaranya adalah “G20 Indonesia 2022”, “KTT G20 Indonesia”, “Pelaksanaan G20 Indonesia”, “Presidensi G20

Indonesia 2022”, dan “Sukseskan KTT G20”. Data yang diperoleh adalah data kotor yang masih diperlukan pembersihan data atau *pre-processing text*.

3.2 Pre-Processing Text

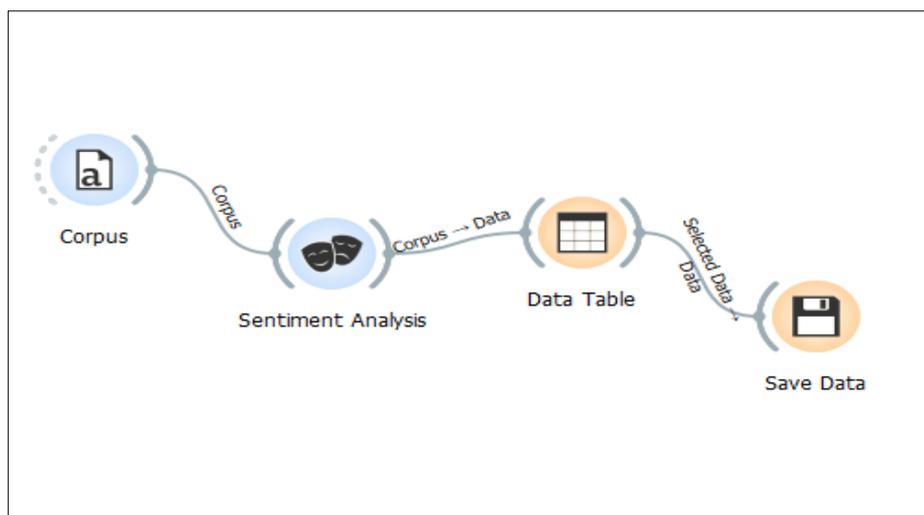
Crawling data dilakukan per-minggu dari bulan Agustus sampai Januari, data tersebut disimpan menjadi satu *file excel* dan mendapatkan 24.840 data kotor atau data awal. Selanjutnya data tersebut akan melalui proses *pre-processing* data, dimana pada proses ini data akan dibersihkan agar bisa lanjut ke tahap pelabelan dan klasifikasi. *Pre-processing* dilakukan dengan menggunakan *software RapidMiner Studio*, berikut adalah alur *Pre-processing text* yang dapat dilihat pada gambar 3.



Gambar 3. Alur *Pre-processing* Teks

3.3 Pemberian Nilai Sentimen dengan *Orange Data Mining*

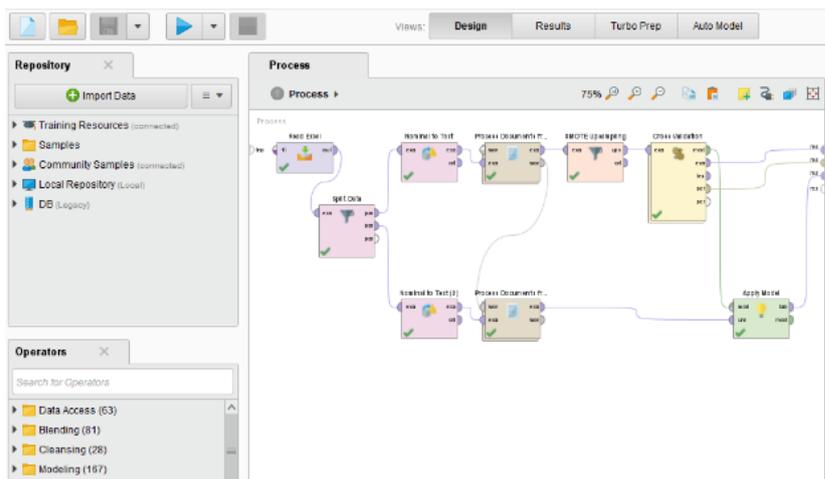
Sebelum dilakukan pelabelan data, nilai sentimen setiap data harus diketahui terlebih dahulu. Untuk menentukan nilai sentimen tersebut dapat menggunakan *software Orange*, dimana dalam *software* ini terdapat beberapa teknik analisis yang dapat digunakan, salah satunya adalah *widget sentiment analysis*. Berikut merupakan alur dari penentuan nilai sentimen dapat dilihat pada gambar 4.



Gambar 4. Alur Sentimen Analisis

3.4 Proses Klasifikasi Data

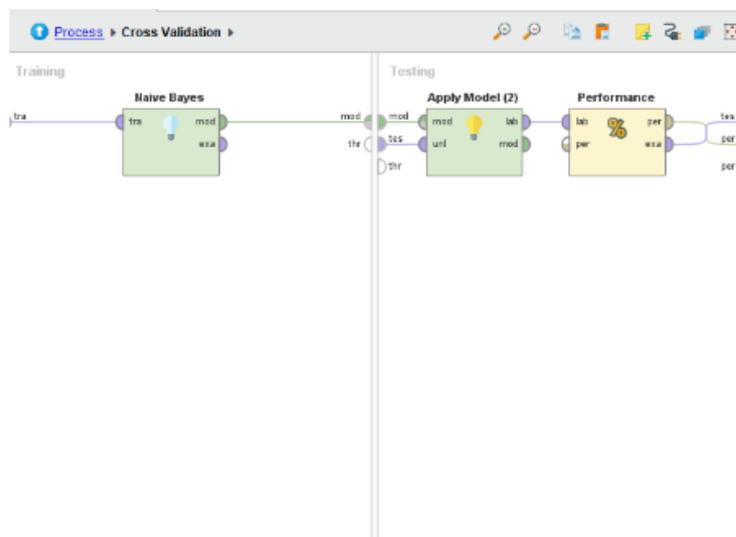
Pada proses klasifikasi data ini akan memaparkan hasil dari sentimen prediksi dan klasifikasi data akan dilakukan dengan menggunakan metode *Naïve Bayes*. Adapun tahapan yang dilakukan sebelum melakukan peng-klasifikasian menggunakan *Naïve Bayes* adalah memasukkan data yang sudah berisi label sentimen kedalam *software RapidMiner*. Kemudian data akan disambungkan dengan operator *split* data untuk membagi dataset menjadi data latih 80% dan data uji 20%. Selanjutnya akan dihubungkan dengan operator-operator lainnya seperti *nominal to text*, *process document form data*, *SMOTE up sampling*, *cross validation*, dan *apply model*. Pada proses klasifikasi ini harus dipastikan untuk tidak adanya tanda *warning* agar saat skema dieksekusi tidak mengalami *error*.



Gambar 5. Skema Dataset dan Operator Klasifikasi Data

3.5 Proses Klasifikasi Naïve Bayes

Pada proses ini merupakan hasil prediksi dari sebuah sentimen akan ditampilkan berdasarkan penerapan dari metode klasifikasi. Berikut ini adalah pengujian dengan metode *naïve bayes* menggunakan skema seperti gambar 6.



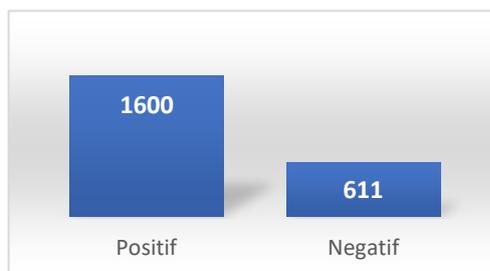
Gambar 6. Skema Metode Naïve Bayes

Pada skema diatas menggunakan operator *cross validation*, yang dimana dalam proses tersebut data latih atau data *training* akan dilatih oleh sistem menggunakan operator *naïve bayes*. Setelah itu menggunakan operator *apply model* untuk menggabungkan antara data *testing* dan data *training* yang sudah dilatih dengan metode *naïve bayes*. Lalu akan dihubungkan dengan operatoor *performance* untuk melihat tingkat *performance* pengujian data menggunakan metode *naïve bayes*. Sehingga metode *naïve bayes* menghasilkan prediksi sentimen sebagai berikut.

Row No.	Sentimen	prediction(Sentimen)	confidence(Positif)	confidence(Negatif)	text
1	Positif	Positif	1	0	sy mendukung presidensi g indonesia piala d
2	Positif	Positif	1	0	salah fiscal tools kementerian keuangan pt pii
3	Positif	Positif	1	0	indonesia tancap gas persiapan sambut ktt g
4	Positif	Positif	1	0	indonesia tancap gas langkahlangkah persiapan
5	Positif	Positif	1	0	sobat bumh menjelang perhelatan ktt g pt wij
6	Negatif	Negatif	0	1	hadapi ancaman serangan siber bnpt jaga kes
7	Positif	Positif	1	0	sobat bumh menjelang perhelatan ktt g pt wij
8	Positif	Positif	1	0	bnpt dukung keamanan siber ktt g indonesia
9	Positif	Positif	1	0	teknologi cisco dilirik ktt g indonesia infrast
10	Positif	Positif	1	0	bocoran harga genesis g electric mobil presid
11	Positif	Positif	1	0	sobat bumh menjelang perhelatan ktt g pt wij
12	Positif	Positif	1	0	menjelang perhelatan ktt g pt wijaya karya pers
13	Positif	Positif	1	0	repost selamat pagi warga bandung tau ga sih
14	Negatif	Positif	1	0	g indonesia ball streaming ball eissas gak kha

Gambar 7. Hasil Prediksi Sentimen Metode Naïve Bayes

Terdapat juga grafik bar dari hasil prediksi sentimen dengan metode *naïve bayes*.



Gambar 8. Grafik Bar Hasil Prediksi Sentimen Naïve Bayes

3.6 Pengujian Akurasi Metode Naïve Bayes

Hasil pengujian metode *naïve bayes* dengan menggunakan *confusion matrix* dapat dilihat pada gambar 9. Pada gambar 9 dijelaskan bahwa pengujian data dengan menggunakan metode *naïve bayes* mendapatkan *total accuracy* 88.01%, dan *precision* dari masing-masing kelas yaitu pred. positif 99.83% dan pred. negatif 80.72%, serta *recall* yang diperoleh pada kelas *true* positif 76.14% dan *true* negatif 99.87%.

accuracy: 88.01% +/- 0.82% (micro average: 88.01%)

	true Positif	true Negatif	class precision
pred. Positif	6044	10	99.83%
pred. Negatif	1894	7928	80.72%
class recall	76.14%	99.87%	

Gambar 9. Confusion Matrix Pengujian Data

3.7 Interpretasi Hasil Klasifikasi

Setelah dilakukan implementasi dan pembahasan pengolahan data sentimen mulai dari pengumpulan data awal, proses pembersihan data atau *pre-processing text* sehingga mendapatkan data bersih, memberikan nilai sentimen dan melakukan pelabelan pada setiap data, serta mengklasifikasi data dengan metode *naïve bayes* sehingga menghasilkan prediksi sentimen terhadap isu kegiatan pelaksanaan G20 tahun 2022. Proses klasifikasi data dengan memanfaatkan *tools-tools* yang terdapat dalam *software RapidMiner Studio* menghasilkan prediksi sentimen yang diuji dengan metode *Naïve Bayes* sebesar total sentimen prediksi positif 1.600 data dan total sentimen prediksi negatif 611 data. Jika dikonversikan kedalam persentase, sentimen positif mendapat 72,37% dan sentimen negatif mendapat 27,63%. Dilihat dari hasil klasifikasi, sentimen yang diberikan oleh masyarakat Twitter cenderung memberikan sentimen positif dibandingkan memberikan sentimen negatif. Hal ini menunjukkan masyarakat sangat mendukung perhelatan G20 tahun 2022 yang diadakan di Bali, Indonesia. Tetapi terdapat juga beberapa masyarakat yang memberikan sentimen negatif. Sehingga dengan mengetahui sentimen yang diberikan masyarakat Twitter terhadap isu kegiatan pelaksanaan G20 tahun 2022 dapat dijadikan sebagai bahan evaluasi atau tolak ukur oleh pemerintah untuk meningkatkan operasional kualitas pengadaan *event* bertaraf internasional yang dilakukan di Indonesia selanjutnya.

Selain menghasilkan total prediksi sentimen, klasifikasi data dengan metode *Naïve Bayes* juga menghasilkan performa *accuracy*, *percission*, dan *recall*, yang bertujuan untuk mengukur performa dari metode *Naïve Bayes* dalam melakukan klasifikasi data sentimen. Proses ini menggunakan perhitungan *confusion matrix*. *Accuracy* yang dihasilkan dalam klasifikasi data dengan *Naïve Bayes* adalah 88.01%. *Precision* yang dihasilkan dari masing-masing kelas yaitu kelas asli positif yang di prediksi benar (positif) oleh sistem sebanyak 6.044 dan kelas asli positif yang diprediksi salah (negatif) oleh sistem sebanyak 10, sehingga mendapat persentase pred. positif 99.83%. Selanjutnya kelas negatif yang diprediksi salah (positif) oleh sistem sebanyak 1.849 dan kelas asli negatif yang diprediksi benar (negatif) oleh sistem sebanyak 7.928, sehingga mendapat persentasi pred. negatif 80.72%. *Recall* yang dihasilkan dari masing-masing kelas yaitu kelas asli positif yang di prediksi benar (positif) sebanyak 6.044 dan kelas asli positif yang di prediksi salah (negatif) sebanyak 1.894 sehingga mendapat persentase 76.14%. Kelas asli negatif yang diprediksi salah (positif) sebanyak 10 dan kelas asli negatif yang diprediksi benar (negatif) sebanyak 7.928 sehingga mendapat persentase 99.87%.

Hasil pengujian model dengan perhitungan *confusion matrix* menunjukkan metode *Naïve Bayes* memiliki performa yang cukup baik dalam topik yang diteliti, dimana didapatkan nilai akurasi sebesar 88.01% keakuratan sistem dapat mengklasifikasi data secara benar. *Precision* pred. positif sebesar 99.83% dan pred. negatif sebesar 80.72% ketepatan sistem antara informasi yang diberikan oleh sistem untuk menunjukan secara benar data kelas negatif atau kelas positif. *Recall* pred. positif sebesar 76.14% dan pred. negatif sebesar 99.87% ketepatan sistem yang menunjukan tingkat keberhasilan atau spesifisitas untuk mengetahui kembali sebuah informasi secara benar tentang data yang kelas negatif atau pun kelas positif.

mengeluarkan, anggaran, recover together recover stronger, tantangan, resesi, salah, ancaman, hoax, siber, dan pemulihan. Sehingga dari kata-kata yang terdapat pada sentimen yang diberikan oleh masyarakat Twitter dapat diartikan sebagai ketidaksetujuan atau keraguan terhadap pelaksanaan G20 tahun 2022 yang dilakukan di Bali, Indonesia sehingga dapat dikatakan sebagai sentimen negatif.

4. KESIMPULAN

Kesimpulan pada penelitian Analisis Sentimen Terhadap Pelaksanaan Kegiatan Presidensi G20 Tahun 2022 Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes yaitu proses klasifikasi dengan metode Naïve Bayes menghasilkan total sentimen prediksi positif 1.600 data dan total sentimen prediksi negatif 611 data. Jika dikonversikan kedalam persentase, sentimen positif mendapat 72,37% dan sentimen negatif mendapat 27,63%, masyarakat Twitter cenderung memberikan sentimen positif dibandingkan memberikan sentimen negatif terhadap isu kegiatan pelaksanaan G20 tahun 2022, Analisis sentimen dilakukan untuk mengetahui sentimen yang diberikan masyarakat Twitter terhadap isu kegiatan pelaksanaan G20 tahun 2022 yang dapat dijadikan sebagai bahan evaluasi atau tolak ukur oleh pemerintah untuk meningkatkan operasional kualitas pengadaan *event* bertaraf internasional yang dilakukan di Indonesia selanjutnya, Hasil pengujian model dengan perhitungan *confusion matrix* menunjukkan metode Naïve Bayes memiliki performa yang cukup baik dalam topik yang diteliti, dimana didapatkan nilai akurasi sebesar 88.01% keakuratan sistem dapat mengklasifikasi data secara benar.

REFERENCES

- [1] A. F. P. I. K. M. I. A. S. Putri, "PENGARUH IMPLEMENTASI LEADERSHIP INDONESIA TERHADAP PRESIDENSI G20 DAN PERTUMBUHAN EKONOMI DI BALI," vol. 8, no. 2, pp. 276–291, 2022.
- [2] N. L. I. Sukawiyana Agus, Anak Agung Elik Astari, "AKTUALISASI PEREKONOMIAN INDONESIA DALAM PRESIDENSI G20 PELUANG BERTUMBUH ATAU TREN? (STUDI ANALISIS EKONOMI)," *Indones. Perspect.*, vol. 7, no. 1, pp. 196–218, 2022, doi: 10.14710/ip.v7i1.48596.
- [3] A. Sasmito Aribowo, "Analisis Sentimen Publik pada Program Kesehatan Masyarakat menggunakan Twitter Opinion Mining," *Semin. Nas. Inform. Medis*, vol. 0, no. 0, pp. 17–23, 2018, [Online]. Available: <https://journal.uui.ac.id/snimed/article/view/11877>
- [4] E. Indrayuni, "Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes," *J. Khatulistiwa Inform.*, vol. 7, no. 1, pp. 29–36, 2019, doi: 10.31294/jki.v7i1.1.
- [5] W. Yulita *et al.*, "Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naive Bayes Classifier," *Jdmsi*, vol. 2, no. 2, pp. 1–9, 2021.
- [6] M. Syarifuddin, "Analisis Sentimen Opini Publik Mengenai Covid-19 Pada Twitter Menggunakan Metode Naive Bayes Dan Knn," *INTI Nusa Mandiri*, vol. 15, no. 1, pp. 23–28, 2020, doi: 10.33480/inti.v15i1.1347.
- [7] R. Tineges, A. Triayudi, and I. D. Sholihati, "Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM)," *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 650, 2020, doi: 10.30865/mib.v4i3.2181.
- [8] A. N. Ulfah and M. K. Anam, "Analisis Sentimen Hate Speech Pada Portal Berita Online Menggunakan Support Vector Machine (SVM)," *JATISI (Jurnal Tek. Inform. dan Sist. Informatika)*, vol. 7, no. 1, pp. 1–10, 2020, doi: 10.35957/jatisi.v7i1.196.
- [9] M. Syarifuddin, "Analisis Sentimen Opini Publik Terhadap Efek Psbb Pada Twitter Dengan Algoritma Decision Tree, Knn, Dan Naive Bayes," *INTI Nusa Mandiri*, vol. 15, no. 1, pp. 87–94, 2020, doi: 10.33480/inti.v15i1.1433.
- [10] D. Musfiroh, U. Khaira, P. E. P. Utomo, and T. Suratno, "Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon: Sentiment Analysis of Online Lectures in Indonesia from Twitter Dataset Using InSet Lexicon," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 24–33, 2021.
- [11] R. W. Samsir, Ambiyar, Unung Verawardina, Firman Edi, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naive Bayes," *J. Media Inform. Budidarma*, vol. 5, no. 1, p. 149, 2021, doi: 10.30865/mib.v5i1.2604.
- [12] D. P. Santoso and W. Wibowo, "Analisis Sentimen Ulasan Aplikasi Buzzbreak Menggunakan Metode Naive Bayes Classifier pada Situs Google Play Store," *J. Sains dan Seni ITS*, vol. 11, no. 2, 2022, doi: 10.12962/j23373520.v11i2.72534.
- [13] E. S. Negara, R. Andryani, and P. H. Saksono, "Analisis Data Twitter: Ekstraksi dan Analisis Data Geospasial," *J. INKOM*, vol. 10, no. 1, p. 27, 2016, doi: 10.14203/j.inkom.433.
- [14] A. Wandani, "Sentimen Analisis Pengguna Twitter pada Event Flash Sale Menggunakan Algoritma K-NN, Random Forest, dan Naive Bayes," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 651–665, 2021.
- [15] D. Alita and A. R. Isnain, "Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier," *J. Komputasi*, vol. 8, no. 2, pp. 50–58, 2020, doi: 10.23960/komputasi.v8i2.2615.
- [16] Normah, B. Rifai, S. Vambudi, and R. Maulana, "Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis SMOTE," *J. Tek. Komput. AMIK BSI*, vol. 8, no. 2, pp. 174–180, 2022, doi: 10.31294/jtk.v4i2.
- [17] W. R. D. Astuti, "Kerja Sama G20 dalam Pemulihan Ekonomi Global dari COVID-19," *Andalas J. Int. Stud.*, vol. 9, no. 2, p. 131, 2020, doi: 10.25077/ajis.9.2.131-148.2020.
- [18] J. Eka Sembodo, E. Budi Setiawan, and Z. Abdurahman Baizal, "Data Crawling Otomatis pada Twitter," no. September, pp. 11–16, 2016, doi: 10.21108/indosc.2016.111.
- [19] A. F. Harismawan, "Analisis Perbandingan Performa Web Service Menggunakan Bahasa Pemrograman Python, Php," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 1, pp. 237–245, 2017, [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/781>
- [20] M. L. Suliztia, "Penerapan Analisis Random Forest pada Prototype Sistem Prediksi Harga Kamera Bekas Menggunakan Flask," *Fak. Mat. Dan Ilmu Pengetah. Alam*, pp. 1–107, 2020.