

Analisis Sentimen Masyarakat Indonesia Terhadap Invasi Russia Di Ukraina Menggunakan Metode *Naïve Bayes* Pada Media Sosial Facebook

Maulana Arvyantomo¹, Niki Ratama²

^{1,2} Teknik, Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia
E-mail: 1maulanaarvy67@gmail.com , 2dosen00835@unpam.ac.id

Abstrak- Invasi Rusia di Ukraina telah membuat media sosial ramai membicarakan hal tersebut terutama pada media sosial Facebook. Sehingga banyak opini - opini atau tanggapan masyarakat Indonesia yang mengomentari dengan lebih mendukung aksi yang dilakukan oleh Rusia dan ada juga yang lebih mendukung aksi yang dilakukan oleh Ukraina sehingga timbul komentar yang bersifat positif dan yang bersifat negatif. Masyarakat Indonesia juga ada yang menengahi atau berkomentar netral atas terjadinya invasi Rusia di Ukraina tersebut. Dalam penelitian ini, penulis akan melakukan analisis sentimen berdasarkan komentar yang membahas tentang invasi Rusia di Ukraina dengan menerapkan metode Naive Bayes untuk proses pengklasifikasian teks apakah termasuk kedalam sentimen positif, negatif maupun netral. Penulis menggunakan 10% dari dataset komentar untuk dijadikan sebagai data latih yang akan digunakan sebagai pembelajaran mesin. Kemudian data diklasifikasi oleh sistem dengan mencari nilai probabilitas tertinggi. Nilai yang lebih tinggi merupakan kelas baru dari data tersebut. Hasil penelitian berupa persentase dari banyaknya sentimen positif, negatif dan netral dari dataset komentar. Metode Naive Bayes terbukti dapat melakukan proses klasifikasi sentimen dengan hasil sebanyak 52% sentimen positif, 5% sentimen negatif dan 43% sentimen netral.

Kata Kunci: Invasi Rusia di Ukraina, Facebook, Komentar, Analisis Sentimen, *Naïve Bayes*

Abstract- *The Russian invasion of Ukraine has made social media busy talking about it, especially on social media Facebook. So that there are many opinions or responses from the Indonesian people who comment more in favor of the actions taken by Russia and some are more supportive of the actions taken by Ukraine so that positive and negative comments arise. There are also Indonesian people who mediate or comment neutrally on the occurrence of the Russian invasion of Ukraine. In this study, the author will conduct a sentiment analysis based on comments that discuss the Russian invasion in Ukraine by applying the Naive Bayes method to the process of classifying texts, whether they are positive, negative or neutral sentiments. The author uses 10% of the comment dataset to serve as training data that will be used for machine learning. Then the data is classified by the system by looking for the highest probability value. The higher value represents the new class of the data. The results of the study are in the form of a percentage of the number of positive, negative and neutral sentiments from the comment dataset. The Naive Bayes method is proven to be able to carry out the sentiment classification process with the results of 52% positive sentiment, 5% negative sentiment and 43% neutral sentiment.*

Keywords: *Russian Invasion in Ukraine, Facebook, Comments, Sentiment Analysis, Naïve Bayes*

1. PENDAHULUAN

Pada tanggal 24 Februari 2022, telah terjadi serangan militer di negara Ukraina. Serangan tersebut dilakukan oleh negara Rusia. Rusia telah melancarkan invasi berskala besar ke Ukraina, dimulai dengan ledakan di sejumlah kota, yakni Kharkiv, Kyiv, Mariupol dan Odessa. Beberapa pejabat dan analisis menyebutkan bahwa invasi ini sebagai serangan militer konvensional terbesar di Eropa sejak Perang Dunia II. Presiden Rusia (Vladimir Putin) menuntut Ukraina dilarang bergabung dengan NATO (North Atlantic Treaty Organization) karena prospek berdirinya NATO di sebelah perbatasan Rusia yang bisa berdampak ancaman bagi keamanan negaranya. Presiden Rusia menyatakan tidak ada rencana untuk menduduki wilayah Ukraina dan ia mendukung hak rakyat Ukraina untuk menentukan nasib sendiri.

Rusia melancarkan serangan besar bertujuan untuk menggulingkan pemerintahan Kyiv melalui cara militer dan memaksa adanya perubahan di Ukraina.

Invasi Rusia di Ukraina mendapat kecaman internasional yang luas, termasuk sanksi - sanksi yang dikenakan pada Rusia. Protes secara global terjadi untuk menentang invasi. Dari konflik tersebut, banyak postingan di media sosial yang membahas tentang invasi Rusia di Ukraina. Salah satu media sosial yang menjadi objek penelitian ini adalah Facebook. Facebook menjadi salah satu media sosial yang banyak digunakan masyarakat Indonesia. Bahkan jumlah pengguna media sosial Facebook ini terbesar nomor 2 di Asia. Berdasarkan data Internetworldstats, pengguna Facebook Indonesia mencapai 175,3 juta pada tahun 2021. Angka tersebut setara dengan 63,4% dari total penduduk yang mencapai 276,36 juta jiwa (estimasi 2021) atau 82% dari pengguna internet di tanah air. Dalam hal ini, banyak opini - opini atau tanggapan masyarakat Indonesia yang bernilai positif, negatif maupun netral. Penulis akan menggunakan perbandingan tersebut terhadap negara Rusia dan Ukraina. Jika tanggapan tersebut lebih dominan mendukung aksi yang dilakukan oleh Rusia maka tanggapan tersebut bernilai positif sedangkan jika tanggapan tersebut lebih dominan mendukung aksi yang dilakukan oleh Ukraina maka tanggapan tersebut bernilai negatif. Apabila tanggapan tidak mendukung aksi yang dilakukan Rusia maupun Ukraina maka tanggapan tersebut bernilai netral.

Untuk menganalisa tanggapan masyarakat Indonesia terhadap invasi Rusia di Ukraina maka dilakukanlah penelitian analisis sentimen. Analisis sentimen merupakan proses mengelompokkan polaritas dari teks yang ada dalam dokumen atau kalimat dan menentukan apakah pendapat tersebut kategorinya bersifat positif, negatif atau netral (Ramadhan & Setiawan, 2019). Dalam melakukan analisis sentimen diperlukan suatu metode. Salah satu metode yang digunakan pada penelitian ini yaitu Naive Bayes. Metode Naive Bayes adalah metode yang bisa memprediksi peluang terjadinya kejadian di masa depan berdasarkan data yang ada sebelumnya (Hakimi, 2018). Kelebihan metode Naive Bayes ini adalah tingkat akurasi yang tinggi dan juga waktu komputasinya yang lebih cepat. Metode ini dirasa cocok digunakan pada analisis sentimen dikarenakan metode ini bertujuan sebagai metode klasifikasi kedalam kategori positif, negatif maupun netral.

Hasil dari analisis sentimen ini bertujuan untuk mengimplementasikan metode Naive Bayes sebagai metode pengklasifikasian analisis sentimen dan mencari nilai persentase dari analisis sentimen pendapat masyarakat Indonesia terhadap invasi Rusia di Ukraina yang bernilai positif, negatif dan netral untuk mengetahui besarnya dampak pengaruh terhadap masyarakat lain.

2. METODE

2.1. Analisis Sentimen

Analisis sentimen merupakan proses mengelompokkan polaritas dari teks yang ada dalam dokumen atau kalimat dan menentukan apakah pendapat tersebut kategorinya bersifat positif, negatif atau netral (Ramadhan & Setiawan, 2019). Saat ini analisis sentimen banyak digunakan oleh peneliti sebagai salah satu cabang riset dalam ilmu komputer seiring dengan ledakan informasi di internet. Ada beberapa penamaan terhadap studi ini, yaitu analisis sentimen, penambangan opini (opinion mining), ekstraksi opini (opinion extraction), penambangan sentimen (sentiment mining) dimana semuanya sekarang berada dibawah ranah analisis sentimen. Analisis sentimen bisa disamakan dengan opinion mining karena berfokus kepada pendapat yang menyatakan positif atau negatif.

Analisis sentimen dapat mengelompokkan teks yang terdapat kalimat, opini, pendapat maupun dokumen. Tujuan dari analisis sentimen untuk menentukan opini dari seseorang dengan memperhatikan topik tertentu (Suryani, Linawati, & Saputra, 2019). Proses analisis bisa mencakup data teks review, forum, tweet, komentar atau blog. Langkah yang dilakukan dalam analisis sentimen adalah crawling, cleaning, case folding, filtering/stopwords, stemming, tokenizing, term frekuensi dan diakhiri dengan proses klasifikasi.

2.2. Masyarakat

Secara umum pengertian masyarakat adalah sekumpulan individu - individu yang hidup bersama, bekerja sama untuk memperoleh kepentingan bersama yang telah memiliki tatanan kehidupan, norma - norma dan adat istiadat yang ditaati dalam lingkungannya (Prasetyo & Irwansyah, 2020). Pengertian masyarakat secara sederhana adalah sekumpulan manusia yang saling berinteraksi atau bergaul dengan kepentingan yang sama. Terbentuknya masyarakat karena manusia menggunakan perasaan, pikiran dan keinginannya memberikan reaksi dalam lingkungannya. Dalam penelitian ini masyarakat menjadi sumber data utama dalam melakukan analisis sentimen, terutama masyarakat yang menggunakan Facebook. Teknologi yang berkembang di masyarakat awal mulanya hanya untuk memudahkan pekerjaan dan komunikasi. Tetapi berjalannya waktu teknologi juga digunakan sebagai saling bertukar pendapat terhadap apa yang sedang terjadi pada saat ini. Banyak pendapat masyarakat mengenai invasi Rusia di Ukraina sehingga penulis mendapat inspirasi untuk membuat judul penelitian terkait invasi tersebut.

2.3. Invasi

Invasi adalah hal atau perbuatan memasuki wilayah negara lain dengan mengarahkan angkatan bersenjata dengan maksud menyerang atau menguasai negara tersebut (KBBI, 2022). Masuknya pasukan Rusia ke wilayah timur Ukraina menandai dimulainya invasi Rusia ke Ukraina. Ketegangan antara Rusia dan Ukraina terjadi setelah Rusia mengakui kemerdekaan dua wilayah timur Ukraina, yaitu Donetsk dan Luhansk yang dikuasai kelompok separatis pro-Rusia pada 21 Februari 2022 (Pakpahan, 2022). Sejak Presiden Putin mengeluarkan instruksi untuk melakukan operasi militer (yang dianggap invasi oleh banyak negara), muncul ketidakpastian ekonomi di berbagai sektor. Ancaman terhadap ketersediaan minyak dan gas bagi negara - negara Eropa. Rusia menguasai 40% pasokan gas alam dan 25% pasokan minyak di Eropa. Hal paling dekat yang akan terjadi adalah berkurangnya pasokan minyak dan gas untuk konsumen Eropa yang masih menghadapi musim dingin dalam setidaknya satu bulan kedepan. Naiknya harga minyak dan gas tentu juga akan berpengaruh kepada aktivitas produksi yang sangat mengandalkan minyak dan gas sebagai bahan baku produksi. Kenaikan harga minyak global tentu akan memberikan pukulan bagi harga jual sektor industri termasuk Indonesia. Dampaknya akan mempengaruhi perekonomian domestik termasuk para produsen lokal Indonesia. Hal ini tentu akan mengganggu proses pemulihan dan pertumbuhan ekonomi Indonesia dan berbagai negara lain.

2.4. Naïve Bayes

Metode *Naïve Bayes* adalah metode yang bisa memprediksi peluang terjadinya kejadian di masa depan berdasarkan data yang ada sebelumnya (Hakimi, 2018). Menurut pendapat lain, *Naïve Bayes* merupakan metode yang menggunakan pendekatan untuk menghasilkan klasifikasi. Metode ini menggabungkan probabilitas term dengan probabilitas kategori untuk menentukan kemungkinan kategori berhasil (Pebrianto, 2019). *Naïve Bayes* akhir-akhir ini banyak digunakan dalam teknik klasifikasi. Kelebihan dari *Naïve Bayes* adalah bahwa metode ini hanya membutuhkan sejumlah kecil data pelatihan untuk memperkirakan parameter yang diperlukan dalam klasifikasi tingkat akurasi yang tinggi dan juga waktu komputasinya yang lebih cepat. Metode ini dirasa cocok digunakan pada analisis sentimen dikarenakan metode ini bertujuan sebagai metode klasifikasi kedalam kategori positif dan negatif. Persamaan yang digunakan dalam metode *Naïve Bayes* adalah sebagai berikut:

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)}$$

Keterangan :

C = Merupakan suatu Class spesifik.

X = Merupakan vektor input (attribut).

P(C|X) = Probabilitas Class berdasarkan input X yang diketahui (posterior probability).

- $P(X|C)$ = Probabilitas tiap input X berdasarkan kondisi pada Class (likelihood).
 $P(C)$ = Probabilitas Class yang dicari (prior probability) dari keseluruhan data.
 $P(X)$ = Probabilitas suatu input dari keseluruhan data.

Dalam klasifikasi Naïve Bayes, kumpulan data yang diproses akan dianotasi dalam tiga kelas, positif, negatif dan netral. Fitur utama klasifikasi Naïve Bayes adalah mendapatkan asumsi yang kuat dari setiap kondisi atau peristiwa (Mesran, et al., 2020). Penulis menggunakan metode ini dalam penulisan penelitian ini.

2.5. Facebook

Facebook merupakan sebuah platform sosial media yang dioperasikan oleh Facebook Inc dan diluncurkan pada tahun 2004. Mark Zuckerberg Bersama 3 orang teman sekamarnya yang telah menciptakan Facebook. Selain teman sekamarnya mereka juga bersama – sama kuliah di jurusan ilmu komputer Harvard. Awalnya untuk penggunaan Facebook dibatasi hanya mahasiswa Harvard saja, kemudian disebarluaskan ke beberapa universitas lain seperti Standford, Ivy League dan Boston karena semakin banyak orang yang tertarik menggunakan Facebook.

Saat ini Facebook adalah media sosial paling besar dalam hal jangkauan global dan jumlah pengguna aktif (Yasya, Muljono, Seminar, & Hardiansyah, 2019). Selain dari jangkauannya yang besar dari segi geografis maupun jumlah, Facebook lebih fokus pada hubungan dan jaringan sosial yang diketahui berpengaruh dalam komunikasi interaktif yang bersifat multimedia dengan kemampuan berbagi teks, gambar, video yang tidak terbatas, serta fitur Group sebagai representasi suatu komunitas atau forum.

2.6. Instant Data Scraper

Instant Data Scraper adalah ekstensi dari browser chrome yang berfungsi sebagai alat mengekstrak data dari halaman web dan mengekspornya sebagai file Excel atau CSV (Webrobots.io, 2022). Tools ini menggunakan AI (Artificial Intelligence) untuk memprediksi data mana yang paling relevan pada halaman HTML dan memungkinkan menyimpannya ke file Excel atau CSV (XLS, XLSX, CSV). Alat ini tidak memerlukan skrip khusus situs web, melainkan menggunakan analisis AI heuristik dari struktur HTML untuk mendeteksi data yang akan diekstraksi. Jika prediksi tidak memuaskan, ini memungkinkan pengguna menyesuaikan pilihan untuk akurasi yang lebih besar. Jenis teknologi pengikisan ini jauh lebih nyaman, karena tidak memerlukan barisan skrip pengikisan yang dibuat oleh pengguna dalam jumlah besar, yang sering kali dipenuhi dengan versi yang sudah ketinggalan zaman dan berlebihan. Selain itu, Instant Data Scraper tidak perlu memiliki keterampilan coding, json, atau xml (Webrobots.io, 2022).

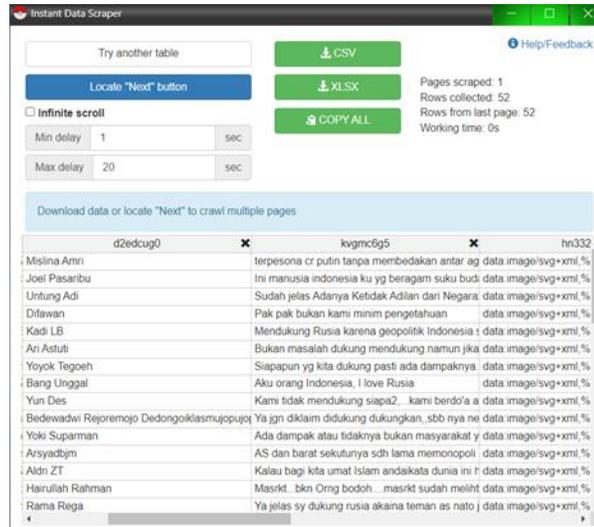
2.7. Jupyter Notebook

Untuk menulis kode program Python dibutuhkan suatu code editor yang disebut Integrated Development Environment atau IDE. Salah satu IDE yang cukup populer terutama di kalangan data scientist untuk menulis kode Python adalah Jupyter Notebook (Aribowo, 2021). Ada beberapa cara dalam melakukan instalasi Jupyter Notebook. Salah satu cara yang paling mudah adalah melalui instalasi Anaconda, yaitu suatu platform yang mendistribusikan bahasa Python khususnya untuk bidang scientific computing (data mining, machine learning, predictive analysis dan lain sebagainya).

3. ANALISA DAN PEMBAHASAN

3.1. *Crawling*

Crawling merupakan proses pengambilan data dari sumber data yang digunakan untuk penelitian. Data berupa komentar diambil dari postingan Facebook Tribunnews Forum yang membahas tentang invasi Rusia di Ukraina. *Crawling* dilakukan menggunakan Instant Data Scraper berikut data hasil *crawling*.



Gambar 1. Data hasil *crawling*

Kemudian dataset disimpan kedalam .csv untuk proses tahap selanjutnya.

3.2. *Labelisasi*

Labelisasi merupakan tahap menentukan sentimen pada data komentar menjadi positif, negatif dan netral kemudian diberi index (nilai) pada setiap sentimen yang bertujuan untuk membedakan setiap nilai sentimen dan dijadikan sebagai data latih untuk penelitian.

	Nama Akun	Sentimen	Komentar
0	Mislina Amri	positif	terpesona cr putin tanpa membedakan antar agam...
1	Rama Rega	positif	Ya jelas sy dukung rusia akaina teman as nato ...
2	Ari Astuti	netral	Bukan masalah dukung mendukung namun jika di d...
3	Dadang Sutsna	netral	Mau negara apapun namanya kalau agres ke nega...
4	Maskur Musholli	netral	as tak berani perang dg jentel slalu melibalka...
5	Wawan Zamali	positif	Saya sangat mendukung Russia karna presiden Pu...
6	Yun Des	netral	Kami tidak mendukung siapa2... kami berdo'a ag...
7	Gandung Widarto	netral	Neraca perdagangan Indonesia dengan Amerika Se...
8	Arnyadbym	positif	AS dan barat sekutunya sdh lama memonopoli asp...
9	Offrex Sticker Ngatrock	positif	Saya bukan tidak peduli sama ukraina karena uk...
10	Caeri Herti	positif	Nampak2 nya ni media uda mulai kagak jelas,men...
11	Ihlin Putra Wijaya	netral	Kalo mendengar cerita orang tua dulu negara ba...

Gambar 2. *Labelisasi*

Data dipilih secara acak sebanyak 50 komentar yakni 10% dari banyaknya dataset hasil *crawling* yang nantinya akan digunakan sebagai data latih klasifikasi *Naïve Bayes*.

3.3. *Cleaning*

Cleaning dilakukan untuk menghapus semua tanda baca pada komentar.

	komentar	hasil_cleaning
0	terpesona cr putin tanpa membedakan antar agama...	terpesona cr putin tanpa membedakan antar agama...
1	Ya jelas sy dukung rusia akaina teman as nato ...	Ya jelas sy dukung rusia akaina teman as nato ...
2	Bukan masalah dukung mendukung namun jika di d...	Bukan masalah dukung mendukung namun jika di d...
3	Mau negara apapun namanya kalau agresi ke nega...	Mau negara apapun namanya kalau agresi ke nega...
4	as tak berani perang dg jentel slalu melibatika...	as tak berani perang dg jentel slalu melibatika...
...
513	Mimin tai media propaganda	Mimin tai media propaganda
514	indonesia harus bersipat netral	indonesia harus bersipat netral
515	Pasti ada dampakx	Pasti ada dampakx
516	Ukraina mau dihadudomba Amirka	Ukraina mau dihadudomba Amirka
517	Iyain aja apa kata tribun	Iyain aja apa kata tribun

518 rows x 2 columns

Gambar 3. Data hasil *cleaning*

3.4. *Case Folding*

Case folding dilakukan untuk menyamakan semua huruf menjadi huruf kecil.

	komentar	hasil_Cleaning	hasil_CaseFolding
0	terpesona cr putin tanpa membedakan antar agama...	terpesona cr putin tanpa membedakan antar agama...	terpesona cr putin tanpa membedakan antar agama...
1	Ya jelas sy dukung rusia akaina teman as nato ...	Ya jelas sy dukung rusia akaina teman as nato ...	ya jelas sy dukung rusia akaina teman as nato ...
2	Bukan masalah dukung mendukung namun jika di d...	Bukan masalah dukung mendukung namun jika di d...	bukan masalah dukung mendukung namun jika di d...
3	Mau negara apapun namanya kalau agresi ke nega...	Mau negara apapun namanya kalau agresi ke nega...	mau negara apapun namanya kalau agresi ke nega...
4	as tak berani perang dg jentel slalu melibatika...	as tak berani perang dg jentel slalu melibatika...	as tak berani perang dg jentel slalu melibatika...
...
513	Mimin tai media propaganda	Mimin tai media propaganda	mimin tai media propaganda
514	indonesia harus bersipat netral	indonesia harus bersipat netral	indonesia harus bersipat netral
515	Pasti ada dampakx	Pasti ada dampakx	pasti ada dampakx
516	Ukraina mau dihadudomba Amirka	Ukraina mau dihadudomba Amirka	ukraina mau dihadudomba amirka
517	Iyain aja apa kata tribun	Iyain aja apa kata tribun	iyain aja apa kata tribun

518 rows x 3 columns

Gambar 4. Data hasil *case folding*

3.5. *Tokenizing*

Tokenizing dilakukan untuk memisahkan setiap kata dan juga dapat menampilkan berapa banyak jumlah setiap kata.

	hasil_Cleaning	hasil_CaseFolding	hasil_Tokenizing
0	terpesona cr putin tanpa membedakan antar agama...	terpesona cr putin tanpa membedakan antar agama...	terpesona, cr, putin, tanpa, membedakan, antar, agama, ...
1	Ya jelas sy dukung rusia akaina teman as nato ...	ya jelas sy dukung rusia akaina teman as nato ...	ya, jelas, sy, dukung, rusia, akaina, teman, as, nato, ...
2	Bukan masalah dukung mendukung namun jika di d...	bukan masalah dukung mendukung namun jika di d...	bukan, masalah, dukung, mendukung, namun, jika, di, d, ...
3	Mau negara apapun namanya kalau agresi ke nega...	mau negara apapun namanya kalau agresi ke nega...	mau, negara, apapun, namanya, kalau, agresi, ke, nega, ...
4	as tak berani perang dg jentel slalu melibatika...	as tak berani perang dg jentel slalu melibatika...	as, tak, berani, perang, dg, jentel, slalu, melibatika, ...
...
513	Mimin tai media propaganda	mimin tai media propaganda	[mimin, tai, media, propaganda]
514	indonesia harus bersipat netral	indonesia harus bersipat netral	[indonesia, harus, bersipat, netral]
515	Pasti ada dampakx	pasti ada dampakx	[pasti, ada, dampakx]
516	Ukraina mau dihadudomba Amirka	ukraina mau dihadudomba amirka	[ukraina, mau, dihadudomba, amirka]
517	Iyain aja apa kata tribun	iyain aja apa kata tribun	[iyain, aja, apa, kata, tribun]

518 rows x 3 columns

Gambar 5. Data hasil *tokenizing*

3.6. *Filtering / Stopwords*

Tahap ini dilakukan untuk memisahkan dan menyaring kata - kata umum yang sering muncul, yang tidak memberikan informasi penting dalam kalimat agar mempermudah proses pengklasifikasian.

	hasil_CaseFolding	hasil_Tokenizing	hasil_Stopwords
0	[terpersona, or, putri, tanpa, membedakan, antar, agam, ...]	[terpersona, or, putri, tanpa, membedakan, antar, agam, ...]	[terpersona, or, putri, membedakan, antar, agam, ...]
1	[ya, jelan, ty, dukung, rusia, ukraina, temen, an, nato, ...]	[ya, jelan, ty, dukung, rusia, ukraina, temen, an, nato, ...]	[jelan, dukung, rusia, ukraina, temen, an, nato, ...]
2	[bukan, masalah, dukung, mendukung, namun, jka, di, d, ...]	[bukan, masalah, dukung, mendukung, namun, jka, di, d, ...]	[masalah, dukung, mendukung, namun, dalam, pa, ...]
3	[mau, negara, apapun, namanya, kalau, agresi, ke, nega, ...]	[mau, negara, apapun, namanya, kalau, agresi, ke, nega, ...]	[negara, namanya, agresi, negara, berdaulat, m, ...]
4	[as, tak, berani, perang, dg, jentel, silau, melibalka, ...]	[as, tak, berani, perang, dg, jentel, silau, m, ...]	[as, berani, perang, jentel, melibalkan, negar, ...]
...
513	[manis, tai, media, propaganda]	[manis, tai, media, propaganda]	[manis, tai, media, propaganda]
514	[indonesia, harus, bersikap, netral]	[indonesia, harus, bersikap, netral]	[indonesia, bersikap, netral]
515	[pauli, ada, dampak]	[pauli, ada, dampak]	[dampak]
516	[ukraina, mau, dituduhdomba, amerika]	[ukraina, mau, dituduhdomba, amerika]	[ukraina, dituduhdomba, amerika]
517	[iyen, apa, apa, kata, tibun]	[iyen, apa, apa, kata, tibun]	[iyen, kata, tibun]

Gambar 6. Data hasil *filtering / stopwords*

3.7. *Stemming*

Stemming dilakukan untuk proses pemetaan dan penguraian berbagai bentuk kata menjadi bentuk dasarnya. Proses pemetaan dan penguraian digunakan untuk menemukan kata dasar dari sebuah kata yang mengalami imbuhan dengan cara menghilangkan atau menghapus imbuhan - imbuhan tersebut.

	hasil_tokenizing	hasil_stopwords	hasil_Stemming
0	[terpersona, or, putri, tanpa, membedakan, antar, agam, ...]	[terpersona, or, putri, membedakan, antar, agam, ...]	[terpersona, or, putri, beda, antar, agama, sing, ...]
1	[ya, jelan, ty, dukung, rusia, ukraina, temen, an, nato, ...]	[jelan, dukung, rusia, ukraina, temen, an, nato, ...]	[jelan, dukung, rusia, ukraina, temen, an, nato, ...]
2	[bukan, masalah, dukung, mendukung, namun, jka, di, d, ...]	[masalah, dukung, mendukung, namun, dalam, pa, ...]	[masalah, dukung, dukung, namun, dalam, perang, ...]
3	[mau, negara, apapun, namanya, kalau, agresi, ke, nega, ...]	[negara, namanya, agresi, negara, berdaulat, m, ...]	[negara, nama, agresi, negara, daulat, turat, ...]
4	[as, tak, berani, perang, dg, jentel, silau, m, ...]	[as, berani, perang, jentel, melibalkan, negar, ...]	[as, berani, perang, jentel, libat, negara, la, ...]
...
513	[manis, tai, media, propaganda]	[manis, tai, media, propaganda]	[manis, tai, media, propaganda]
514	[indonesia, harus, bersikap, netral]	[indonesia, bersikap, netral]	[indonesia, sipat, netral]
515	[pauli, ada, dampak]	[dampak]	[dampak]
516	[ukraina, mau, dituduhdomba, amerika]	[ukraina, dituduhdomba, amerika]	[ukraina, dituduhdomba, amerika]
517	[iyen, apa, apa, kata, tibun]	[iyen, kata, tibun]	[iyen, kata, tibun]

Gambar 7. Data hasil *stemming*

3.8. *Term Frequency*

Untuk Pada penelitian ini pembobotan *term frequency* dilakukan dengan cara mencari jumlah kemunculan sebuah kata dalam tiap dokumen dari sekumpulan data *training*.

TF_Positif	TF_Negatif	TF_Neutral
[('rusia', 18), ('dukung', 14), ('negara', 14), ('as', 11), ('indonesia', 9), ('putin', 8), ('nato', 7), ('tidak', 7), ('ukraina', 5), ('barat', 5), ('russia', 4), ('media', 4), ('masyarakat', 4), ('tdk', 4), ('lindung', 3), ('jelas', 3), ('palistina', 3), ('rakyat', 3), ('salah', 3), ('dunia', 3), ...]	[('negara', 6), ('ukraina', 4), ('pikir', 3), ('nato', 3), ('dukung', 2), ('rusia', 2), ('jadi', 2), ('anggota', 2), ('kuat', 2), ('wel', 1), ('diri', 1), ('dikit', 1), ('ngapain', 1), ('indonesia', 1), ('maju', 1), ('minyak', 1), ('goreng', 1), ('mudah', 1), ('atau', 1), ('jauh', 1), ...]	[('negara', 20), ('as', 12), ('dukung', 11), ('ukraina', 9), ('lain', 9), ('rusia', 8), ('indonesia', 7), ('perang', 6), ('tidak', 6), ('dampak', 6), ('rakyat', 5), ('damai', 5), ('jadi', 4), ('banyak', 4), ('lihat', 3), ('invasi', 3), ('amerika', 3), ('sekutu', 3), ('serang', 3), ('nato', 3), ...]

Gambar 8. TF Positif, TF Negatif dan TF Netral

3.9. Term Frequency

a. Klasifikasi Data Latih

1. Menghitung Probabilitas Prior Setiap Kategori “P(C)”

Rumus :

$$P(C) = \frac{\sum TF|(Positif/Negatif/Netral)}{\sum TF}$$

$\sum TF|(Positif/Negatif/Netral)$ = Jumlah term frekuensi setiap kategori.

$\sum TF$ = Jumlah dari keseluruhan term frekuensi ($TF_{positif} + TF_{negatif} + TF_{netral}$).

Berikut adalah hasil yang diperoleh dari perhitungan probabilitas prior.

G	H	I	J	K
*P(C)	Positif	Negatif	Netral	Total
P(Positif/Negatif/Netral)	0.48855	0.074155	0.437296	100%

Gambar 9. Probabilitas *prior*

2. Menghitung Probabilitas Likelihood Setiap Term “P(X|C)”

Rumus :

$$P\left(X \left| \begin{array}{l} \text{Positif} \\ \text{Negatif} \\ \text{Netral} \end{array} \right.\right) = \frac{nTF(Positif/Negatif/Netral) + 1}{\sum TF|(Positif/Negatif/Netral) + |r|}$$

nTF = Banyak nya kata pada setiap kategori.

$\sum TF|(Positif/Negatif/Netral)$ = Jumlah term frekuensi setiap kategori.

$|r|$ = Jumlah banyaknya kata.

Berikut adalah beberapa data hasil yang diperoleh dari perhitungan probabilitas *likelihood*.

488		448	68	401	917	100%	100%	100%
No	Kata	TF Positif	TF Negatif	TF Netral	Jumlah	P(Positif)	P(Negatif)	P(Neutral)
1	rusia	18	2	8	28	0.02030	0.00540	0.01012
2	dukung	14	2	11	27	0.01603	0.00540	0.01350
3	negara	14	6	20	40	0.01603	0.01259	0.02362
4	as	11		12	23	0.01282	0.00180	0.01462
5	indonesia	9	1	7	17	0.01068	0.00360	0.00900
6	putin	8	1	1	10	0.00962	0.00360	0.00225
7	nato	7	3	3	13	0.00855	0.00719	0.00450
8	tidak	7		6	13	0.00855	0.00180	0.00787
9	ukraina	5	4	9	18	0.00641	0.00899	0.01125
10	barat	5		2	7	0.00641	0.00180	0.00337
11	russia	4			4	0.00534	0.00180	0.00112
12	media	4		1	5	0.00534	0.00180	0.00225
13	masyarakat	4			4	0.00534	0.00180	0.00112
14	tdk	4		1	5	0.00534	0.00180	0.00225
15	lindung	3			3	0.00427	0.00180	0.00112
16	jelas	3		1	4	0.00427	0.00180	0.00225
17	palistina	3			3	0.00427	0.00180	0.00112
18	rakyat	3	1	5	9	0.00427	0.00360	0.00675
19	salah	3		1	4	0.00427	0.00180	0.00225
20	dunia	3		2	5	0.00427	0.00180	0.00337
21	timur	3		1	4	0.00427	0.00180	0.00225
22	ukraena	3			3	0.00427	0.00180	0.00112
23	baik	3		1	4	0.00427	0.00180	0.00225
24	nyh	3			3	0.00427	0.00180	0.00112
25	untung	3			3	0.00427	0.00180	0.00112
26	allah	2	1		3	0.00321	0.00360	0.00112
27	anak	2			2	0.00321	0.00180	0.00112
28	biar	2			2	0.00321	0.00180	0.00112
29	masuk	2		1	3	0.00321	0.00180	0.00225
30	israel	2		2	4	0.00321	0.00180	0.00337
31	buat	2			2	0.00321	0.00180	0.00112
32	sekutu	2		3	5	0.00321	0.00180	0.00450
33	sekarang	2		1	3	0.00321	0.00180	0.00225
34	maju	2	1	2	5	0.00321	0.00360	0.00337
35	mulai	2		1	3	0.00321	0.00180	0.00225
36	boneka	2			2	0.00321	0.00180	0.00112
37	moga	2		1	3	0.00321	0.00180	0.00225
38	pak	2		3	5	0.00321	0.00180	0.00450
39	uraaa	2			2	0.00321	0.00180	0.00112
40	percaya	2			2	0.00321	0.00180	0.00112
41	buruk	2			2	0.00321	0.00180	0.00112
42	indo	2			2	0.00321	0.00180	0.00112
43	berat	2		1	3	0.00321	0.00180	0.00225

Gambar 10. Probabilitas *likelihood*

b. **Klasifikasi Data Uji**

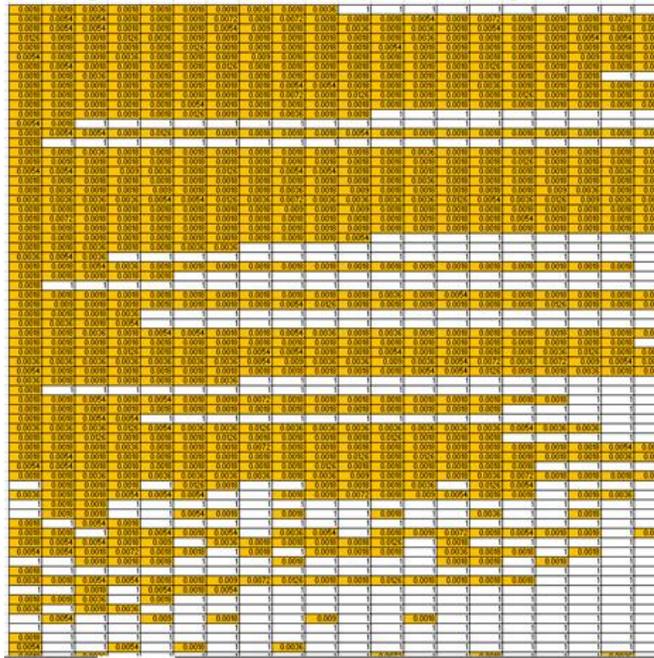
Proses klasifikasi data uji dengan mengalikan semua nilai peluang yang ada dalam kelas. Nilai yang lebih tinggi merupakan kelas baru dari data tersebut.

Rumus :

$$P(C|X) = P(X_1|C) \cdot P(X_2|C) \dots P(X_n|C) \cdot P(C)$$

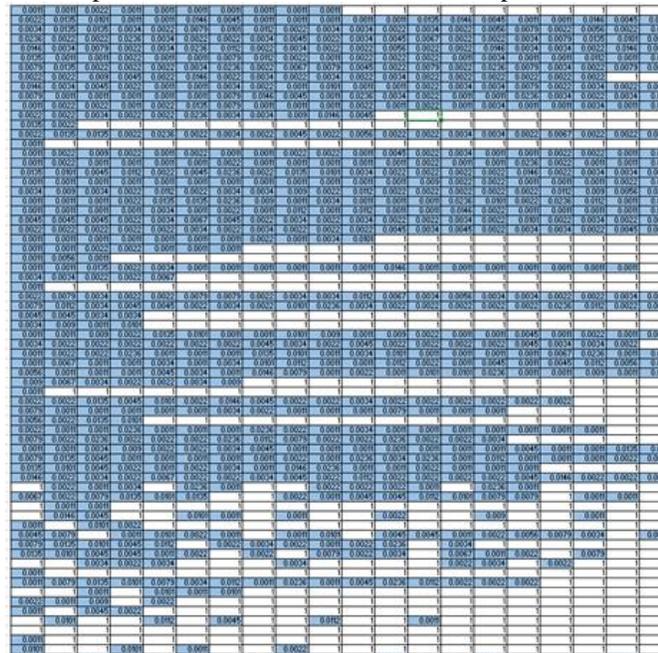
Berikut merupakan sebagian dari tampilan keseluruhan data berupa kata - kata yang telah termasuk kedalam data latih.

Berikut merupakan nilai probabilitas *likelihood* negatif dari setiap *term*.



Gambar 13. Data uji menggunakan probabilitas *likelihood* negatif

Berikut merupakan nilai probabilitas *likelihood* netral dari setiap *term*.



Gambar 14. Data uji menggunakan probabilitas *likelihood* netral



REFERENSI

- Aribowo, B. (2021). Belajar Python Dari Nol Via Jupyter Notebook. Yogyakarta: Diandra Kreatif.
- Hakimi, F. D. (2018). Sistem Analisis Sentimen Publik Tentang Opini Pemilihan Kepala Daerah Jawa Timur 2018 Pada Dokumen Twitter Menggunakan Naive Bayes Classifier. Skripsi, Surabaya : Universitas Islam Negeri Sunan Ampel Surabaya.
- KBBI. (2022, Juni 08). KBBI Daring. Diambil kembali dari kbbi.kemdikbud.go.id/entri/invasi
- Mesran, Sulaiman, O. K., Wijoyo, H., Putra, S. H., Watrianthos, R., Sinaga, R., . . . Indarto, S. L. (2020). Merdeka Kreatif di Era Pandemi Covid-19. Medan: Green Press.
- Pakpahan, A. K. (2022, Februari 24). Invasi Rusia Ke Ukraina Dan Perekonomian Global. unpar.ac.id. Diambil kembali dari <https://unpar.ac.id/invasi-rusia-ke-ukraina-dan-perekonomian-global/>
- Pebrianto, J. (2019). Analisis Sentimen Layanan Internet Provider Selular Pada Tweet Twitter Menggunakan Metode Naive Bayes Classifier Dengan Bahasa Pemrograman PHP. Skripsi, Jakarta : STMIK ERESHA.
- Prasetyo, D., & Irwansyah. (2020, Januari). Memahami Masyarakat Dan Perspektifnya. JMPIS, Vol 1(1), 163.
- Ramadhan, D. A., & Setiawan, E. B. (2019, Agustus). Analisis Sentimen Program Acara Di SCTV Pada Twitter Menggunakan Metode Naive Bayes Dan Support Vector Machine. e-Proceeding of Engineering, 6, 9737.
- Suryani, N. P., Linawati, & Saputra, K. O. (2019, 04 15). Penggunaan Metode Naive Bayes Classifier pada Analisis Sentimen Facebook Berbahasa Indonesia. Majalah Ilmiah Teknologi Elektro, 18 No. 1, 145.
- Webrobots.io. (2022, Mei 26). <https://chrome.google.com/webstore/detail/instant-data-scraper/ofaokhiedipichpaobibbnahnkdoiiah>, 1.0.7. Dipetik Mei 26, 2022, dari <https://webrobots.io/instantdata>.
- Yasya, W., Muljono, P., Seminar, K. B., & Hardiansyah. (2019, Mei). Pengaruh Penggunaan Media Sosial Facebook Dan Dukungan Sosial Online Terhadap Perilaku Pemberian ASI. Studi Komunikasi Dan Media, 72.