

## Analisis Sentimen Tweet Terhadap Penggunaan Layanan JNE & JNT Menggunakan Metode Naive Bayes

Clarisa Novita<sup>1\*</sup>, Farida Nurlaila<sup>1</sup>

<sup>1</sup>Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Jl. Raya Puspiptek No. 46, Kel. Buaran, Kec. Serpong, Kota Tangerang Selatan. Banten 15310, Indonesia

Email: [1\\*Clareisacn.08@gmail.com](mailto:1*Clareisacn.08@gmail.com), [2faridanurlaila09@gmail.com](mailto:2faridanurlaila09@gmail.com)

(\* : coressponding author)

**Abstrak**— Jasa pengiriman saat ini merupakan hal yang dekat dengan masyarakat, jasa pengiriman masih digunakan masyarakat karena dirasa sangat penting untuk mengirim barang dengan jarak yang jauh ataupun dekat. Karena sebagaimana yang telah kita ketahui bersama, Sebagaimana yang telah kita ketahui bersama, bahwa dalam rangka percepatan penanganan Covid-19 (Corona Virus Disease 2019), pemerintah mengeluarkan PP Nomor 21 Tahun 2020 tentang PSBB (Pembatasan Sosial Berskala Besar). Dengan diberlakukannya PSBB membawa budaya baru pada digitalisasi di banyak sektor, termasuk sektor publik sosial. Sehingga sebagian besar kegiatan berbelanja masyarakat terpaksa harus beralih dengan menggunakan sistem online dan membuat jasa pengiriman kini sangat berfungsi pada saat ini. Maka dari itu peneliti bermaksud untuk melakukan penelitian ini yang ditujukan supaya mengetahui kepuasan pelanggan terhadap Ekspedisi pengiriman JNE dan JNT. Penelitian dilakukan dengan melakukan pengambilan sentiment tweet menggunakan Python, lalu kemudian dilakukan analisis sentiment responden menggunakan Metode Naive Bayes. Sebanyak 955 data didapat, lalu peneliti mengambil sebanyak 243 untuk diuji kembali, pengambilan data dilakukan pada tanggal 31 Januari 2022 sampai dengan 16 February 2022. Hasil dari klasifikasi tweet sentiment positif, negative dan netral dari kalimat yang telah ada untuk masing masing jasa pengiriman telah diperoleh bahwa JNE mendapat 0 respon positif, 18 respon negative dan 56 respon netral. Sedangkan JNT mendapat 0 respon positif, 2 respon netral dan 167 respon negative. Dari hal ini dapat dijelaskan bahwa masih banyak pengguna layanan ekspedisi pengiriman JNE dan JNT belum mendapatkan pelayanan dengan baik. Hasil dari pengujian yang menggunakan metode naive Bayes dengan rumus Densitas Gauss pada data kontinu menghasilkan akurasi 58.02 %, presisi 43.33%, recall 41.67%.

**Kata Kunci:** Ekspedisi Pengiriman, Naive Bayes, Python, Scraping

**Abstract**— Delivery service is something that is close to the community, delivery services are still used by the community because it is considered very important to send goods over long or close distances. Because as we all know, as we all know, that in the context of accelerating the handling of Covid-19 (Corona Virus Disease 2019), the government issued PP No. 21 of 2020 concerning PSBB (Large-Scale Social Restrictions). With the enactment of PSBB, it brings a new culture of digitization in many sectors, including the social public sector. So that most people's shopping activities are forced to switch to using an online system and make delivery services very functional at this time. Therefore, the researcher intends to conduct this research which is intended to determine customer satisfaction with JNE and JNT shipping expeditions. The research was conducted by taking tweet sentiment using Python, and then analyzing the respondents' sentiment using the Naive Bayes method. A total of 955 data were obtained, then the researchers took as many as 243 for retesting, data collection was carried out on January 31, 2022 to February 16, 2022. The results of the classification of positive, negative and neutral sentiment tweets from existing sentences for each delivery service have been obtained that JNE got 0 positive responses, 18 negative responses and 56 neutral responses. Meanwhile, JNT received 0 positive responses, 2 neutral responses and 167 negative responses. From this it can be explained that there are still many users of JNE and JNT shipping expedition services that have not received good service. The results of the test using the Nave Bayes method with the Gaussian Density formula on continuous data resulted in an accuracy of 58.02%, precision 43.33%, recall 41.67%.

**Keywords:** Shipping Expedition, Nave Bayes, Python, Scraping

## 1. PENDAHULUAN

Jasa pengiriman saat ini merupakan hal yang dekat dengan masyarakat, jasa pengiriman masih digunakan masyarakat karena dirasa sangat penting untuk mengirim barang dengan jarak yang jauh ataupun dekat. Karena sebagaimana yang telah kita ketahui bersama, Sebagaimana yang telah kita ketahui bersama, bahwa dalam rangka percepatan penanganan Covid-19 (Corona Virus Disease 2019), pemerintah mengeluarkan PP Nomor 21 Tahun 2020 tentang PSBB (Pembatasan Sosial Berskala Besar). Dengan diberlakukannya PSBB membawa budaya baru pada digitalisasi di banyak sektor, termasuk sektor publik sosial. Sehingga sebagian besar kegiatan berbelanja masyarakat terpaksa harus beralih dengan menggunakan sistem online dan membuat jasa pengiriman kini sangat berfungsi pada saat ini.

Berdasarkan data yang dikeluarkan oleh (APJII, 2020) (Asosiasi Penyelenggara Jasa Internet Indonesia) jumlah pengguna internet di Indonesia hingga kuartal II tahun 2020 adalah 196,7 juta orang, naik menjadi 73,7% dari populasi. Hal ini membuat internet memegang peran yang cukup krusial selama WFH (*Work From Home*) dan Pembelajaran Jarak Jauh, karena banyak kegiatan yang kemudian dialihkan menjadi *online* yang mana memaksa seluruh masyarakat tetap beraktivitas dengan normal meski tidak bertatap muka. Hal ini menunjukkan digitalisasi di masa pandemi terasa begitu masif. Sejalan dengan itu, kebutuhan akan layanan internet pun melonjak. Penyedia layanan internet atau *Internet Service Provider* (ISP) pun merasakan dampaknya karena kebutuhan akses internet yang tinggi berbanding lurus dengan pertumbuhan penyedia layanan internet. Sehingga layanan *fixed broadband* menjadi daya tarik bagi beberapa orang untuk mempermudah pelaksanaan aktivitasnya di masa pandemi.

Dengan beralihnya kebutuhan berbelanja menggunakan jasa online, maka peran jasa pengiriman sangatlah penting, diperlukan kecerdasan dalam memilih layanan jasa pengiriman. Terutama untuk memilih jasa apa yang memiliki kualitas layanan yang handal dan terpercaya. Salah satu faktor yang dapat dijadikan tolak ukur dalam memilih jasa pengiriman adalah kualitas pada pengiriman. Dimana kualitas sebuah pengiriman pada suatu perusahaan jasa dipengaruhi dengan tingkat kepuasan pelanggan, karena umumnya kepuasan pelanggan menjadi bahan evaluasi bagi penyedia jasa untuk memperbaiki dan meningkatkan kualitas pelayanannya.

Oleh sebab itu, penelitian ini akan mencoba untuk mengukur tingkat kepuasan pelanggan dan menganalisis serta membandingkan dari berbagai opini atau sentimen pengguna terhadap layanan terhadap kedua jasa pengiriman pada jejaring sosial Twitter. Jasa pengiriman yang dipilih pada penelitian ini yaitu JNE & JNT, karena berdasarkan survey Frekuensi penggunaan jasa kurir jika dibandingkan dengan sebelum pandemi diakui oleh 39% responden meningkat signifikan dan 39% lainnya mengaku sedikit meningkat. Mayoritas masyarakat menggunakan jasa kurir untuk mengirimkan barang yang dibeli dari e-commerce sebanyak 85,2%. Sedangkan sebesar 50,8% lainnya mengirimkan barang belanjaan dari tempat belanja online di media sosial. Di masa pandemi, masyarakat menjadi semakin selektif dalam memilih jasa ekspedisi yaitu sebesar 72%.

Twitter merupakan jejaring sosial dengan fitur *micro-blogging* yang memiliki jumlah pengguna dan *database* yang besar dan terus berkembang. Sering kali twitter dijadikan tempat untuk menyampaikan opini karena memungkinkan penggunaannya untuk berbagi pesan menggunakan teks pendek yang disebut *tweet*. Dengan banyaknya opini yang muncul, dapat dimanfaatkan menjadi landasan untuk mengetahui sentimen terhadap jasa atau produk. Serta berguna untuk mengetahui tingkat kepuasan pelanggan terhadap suatu ekspedisi, sehingga dapat dijadikan bahan evaluasi terhadap layanan JNE & JNT dalam meningkatkan kualitas pelayanannya.

Menurut (Liu, 2011) *sentiment analysis* atau *opinion mining* mengacu pada bidang yang luas dari pengolahan bahasa alami, komputasi linguistik dan teks mining yang memiliki tujuan menganalisa pendapat, sentimen, evaluasi, sikap, penilaian dan emosi seseorang apakah pembicara atau penulis berkenan dengan suatu topik, produk, layanan, organisasi, individu, ataupun kegiatan tertentu. Algoritma yang digunakan dalam penerapan metode analisis sentimen terhadap layanan JNE & JNT yaitu *naive bayes*, berdasarkan penelitian Rish (Mustofa Hidayat & Syafrullah, 2017) menunjukkan bahwa meskipun asumsi independensi antar kata dalam dokumen tidak sepenuhnya dapat dipenuhi, tetapi kinerja *Naive Bayes Classifier* dalam klasifikasi *relative* sangat bagus. Dari kesimpulan pengujian (Sudiantoro & Zuliarso, 2018) Algoritma *Naive Bayes Classifier* sangat efektif untuk digunakan sebagai proses klasifikasi *tweet* yang dibutuhkan dalam sistem analisis *sentiment* ini dimana nilai yang di dapatkan dalam pengujian sampai 84%.

Berdasarkan latar belakang yang telah dijelaskan diatas, maka dapat diambil beberapa masalah yang akan dibahas dalam penulisan ini yaitu seberapa tinggi tingkat kepuasan pelanggan pada jasa pengiriman JNE & JNT di jejaring media sosial Twitter. Dan mengukur tingkat akurasi yang didapatkan dengan menggunakan metode *naive bayes*. Adapun masalah yang menjadi batasan pekerjaan dalam penulisan Tugas akhir ini antara lain, pengumpulan data melalui Tweets yang dilakukan dengan cara scrapping dari twitter, berdasarkan kata kunci “JNE” dan “JNT” dan respon masyarakat dari kedua penyedia jasa pengiriman tersebut. Data sentimen akan diklasifikasikan dalam dua kelas yaitu pendapat positif dan negatif, pencarian tweets dilakukan secara manual dan otomatis dengan pendekatan kamus sentimen.

Dengan ini diharapkan hasil yang didapat dari pembuatan aplikasi dapat membantu penyedia layanan pengiriman dalam meningkatkan kualitas layanan dengan melakukan evaluasi dan juga untuk mempermudah calon pengguna layanan pengiriman dalam memilih jasa yang akan digunakan.

## 2. METODOLOGI PENELITIAN

Hasil penelitian ini diharapkan dapat bermanfaat dalam beberapa aspek sebagai berikut :

- a. Hasil analisis ini nantinya dapat digunakan sebagai tolak ukur dalam melakukan evaluasi pada strategi bisnis serta diharapkan bisa meningkatkan kualitas pelayanan bagi para jasa pengiriman, berdasarkan opini ataupun sentimen pelanggan terhadap penyedia jasa tersebut.
- b. Hasil analisis ini dapat digunakan untuk calon pengguna sebagai bahan pertimbangan dalam menentukan jasa pengiriman mana yang akan digunakan.

Pada bagian ini akan diuraikan tentang beberapa penelitian yang sudah dilakukan sebelumnya. Berikut ini adalah uraian dari beberapa penelitian yang berkaitan dengan topik penelitian yang penulis lakukan:

1. “ANALISIS SENTIMEN PELANGGAN TOKO ONLINE JD.ID MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER BERBASIS KONVERSI IKON EMOSI” (Fransiska vina sari, Volume 10 – ISSN : 2549.3108, 2019) Analisis Sentimen adalah suatu teknik mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif, netral maupun negatif. Analisis sentimen diberikan oleh pengguna internet pada media sosial untuk memberikan suatu penilaian atau opini pribadi. Salah satu toko online Indonesia yang sering mendapatkan sentimen pengguna melalui media sosial adalah JD.id. Adanya sentimen opini dari konsumen tentang JD.id dapat dianalisis dan dimanfaatkan untuk mendapatkan informasi yang berguna bagi pelanggan lain maupun pihak toko. Dengan menggunakan teknik *Text Mining* metode klasifikasi, akan diketahui suatu sentimen bernilai positif, netral atau negatif. Salah satu algoritme yang banyak digunakan dalam analisis sentimen adalah metode klasifikasi *Naive Bayes*. Penelitian ini menggunakan metode *Naive Bayes Classifier (NBC)* dengan pembobotan *tf-idf* disertai penambahan fitur konversi ikon emosi(emoji) untuk mengetahui kelas sentimen yang ada dari *tweet* tentang toko JD.id. Hasil penelitian menunjukkan bahwa metode *Naive Bayes* tanpa penambahan fitur mampu mengklasifikasi sentimen dengan nilai akurasi sebesar 96,44%, sementara jika ditambahkan fitur pembobotan *tf-idf* disertai konversi ikon emosi mampu meningkatkan nilai akurasi menjadi 98%.
2. “ANALISIS SENTIMEN TWITTER UNTUK REVIEW FILM MENGGUNAKAN ANALISIS NAIVE BAYES CLASIFFIER (NBC) PADA SENTIMEN PROGRAMMING” (Fransiska, volume 5 – ISSN 2615 – 7465, 2019) Twitter merupakan salah satu media sosial yang memungkinkan penggunaannya bebas berkomentar atau menulis tweet opini tentang apapun yang terjadi, termasuk opini tentang film. Banyaknya tweet komentar atau opini yang ditulis di twitter dapat dimanfaatkan untuk mendapatkan informasi dan penilaian film yang akan menjadi referensi bagi pengguna twitter lain. Akan tetapi dengan banyaknya tweet membuat kesulitan bagi pengguna yang ingin melihat opini tentang review film tersebut karena data tweet masih terusun secara acak. Agar diketahui opini secara tersusun sesuai dengan klasifikasi maka dalam penelitian ini dilakukan analisis sentimen dengan

metode Naive Bayes. Judul film yang menjadi fokus dalam penelitian ini adalah "Parasite". Data yang dihasilkan dari tahap analisis sentiment ini dikelompokkan menjadi dua jenis sentimen. Pertama sentiment emotion dengan mengklasifikasikan data menjadi enam: emotion, anger, disgust, fear, joy, sadness, surprise. Kedua sentiment polarity mengklasifikasikan data menjadi sentimen positif, sentimen negatif dan sentimen netral. Dengan dua klasifikasi yang berbeda didapatkan hasil data dari sentiment emotion lebih banyak unknown dan untuk sentiment polarity lebih banyak sentimen negatif. Ukuran data dari data awal yang sebanyak 99.521 tweet dengan ukuran 13.9 Mb menjadi 7.5 MB dan data tweet menjadi sebanyak 36.036 tweet yang telah diperoleh.

3. "ANALISIS SENTIMEN DOKUMEN TWITTER MENGENAI DAMPAK VIRUS CORONA MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER" (Ni Made Ayu Juli Astari. Jurnal On jurnal sistem dan informatika, Volume 15 No. 1 – ISSN: 2460 – 3732, 2020) Virus Corona menjadi permasalahan internasional pada tahun 2020. Hal ini sangat berdampak bagi kehidupan masyarakat. Pemerintah Indonesia mengambil peran dalam menekan peningkatan jumlah penderita virus Corona dengan cara membatasi kegiatan masyarakat di luar rumah. Salah satu dampak yang signifikan dari Virus Corona adalah di sektor perekonomian. Oleh karena itu, perlu dilakukan analisis sentimen untuk menentukan kecenderungan opini masyarakat terhadap dampak virus Corona. Twitter merupakan salah satu platform yang digunakan oleh masyarakat untuk mengekspresikan kondisi terkini setelah virus Corona merambah. Tujuan dari penelitian ini adalah memperoleh analisis dokumen text untuk mendapatkan sentimen positif atau negatif masyarakat. Data yang digunakan merupakan dokumen tweet dari Twitter mengenai dampak virus Corona. Data yang terkumpul dibagi untuk digunakan sebagai data latih dan data uji proses klasifikasi. Metode yang digunakan untuk klasifikasi dalam penelitian ini adalah Metode Naive Bayes Classifier. Hasil klasifikasi dievaluasi menggunakan accuracy dan error rate dengan tujuan mengetahui keakuratan dokumen setelah diklasifikasi menjadi sentimen positif atau negatif. Hasil penelitian menunjukkan metode Naive Bayes mampu mengklasifikasi dokumen tweet dengan akurasi 67% dan error rate sebesar 33%. Percobaan dengan menggunakan 3 jumlah data berbeda (100, 200, dan 500) menghasilkan selisih nilai akurasi yang tidak jauh berbeda yaitu 0,02. Hal ini menunjukkan metode Naive Bayes untuk klasifikasi data tweet terkait dampak virus Corona menghasilkan performa yang stabil. Nilai accuracy yang diperoleh cukup baik dan penelitian selanjutnya bisa dikembangkan dengan memperhitungkan unsur semantik pada dokumen tweet.

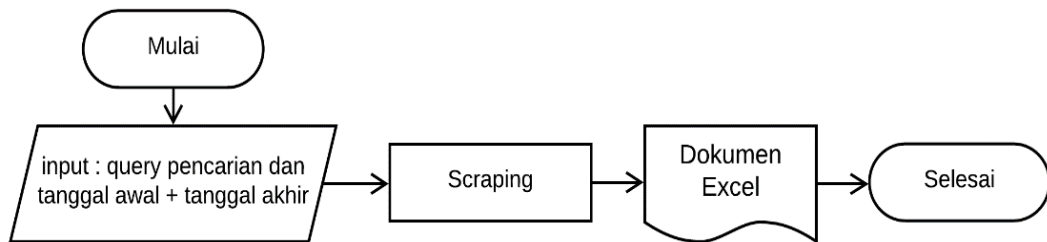
### 3. ANALISA DAN PEMBAHASAN

Data yang digunakan merupakan kumpulan data set yang telah dirangkum pada penelitian ini, data yang diambil bersumber dari jejaring sosial twitter yang telah dikumpulkan mulai tanggal 1 September 2021 sampai nanti selesai pada tanggal 30 September 2021, yang diambil pada masa PPKM level 3 pulau Jawa Bali, data tersebut diambil menggunakan sistem scrapping dibangun dengan bahasa pemrograman Python dengan memanfaatkan Twitter API menggunakan library Tweepy. Data set dibagi menjadi 2 secara seimbang untuk masing-masing provider, dan dibagi 2 untuk data training dan testing pada masing-masing provider dengan rasio 80:20. Data yang dikumpulkan merupakan data teks berupa tweet berbahasa Indonesia dari masyarakat berdasarkan keywords: JNT DAN JNE. Kemudian data tweet disimpan dalam file xlsx (Excel). Berikut adalah beberapa record yang telah diambil dari tweet hasil scraping dapat dilihat pada Tabel 1 berikut:

**Tabel 1.** Hasil Data Scraping

<i>Username</i>	<i>Tweet</i>	<i>Date</i>	<i>Category</i>
cheesycheez	udah datengnya malam, paket dilempar gitu aja kurir ga manggil atau apapun, emang JNE paling bagus deh	2021-09-26 19.28	JNE
bjocah	thanks JNE!	2021-09-26 19.26	JNE
kembangmlathi	Halo @JNECare status paket saya sudah delivered ke saya, sejak kemarin malam. Tapi saya belum terima sampai sekarang. Bagaimana ini ? Resi : JT36390901902	2021-09-26 18.42	JNE
mimiaprl	jangan dulu update, rson nih komplainan saya. Udah mau sebulan paket ga nyampe2 gatau dimana, di telpon jawabnya tunggu2 mulu	2021-09-27 13.51	JNT
rifaldiabdul81	min resi ini udah delivered tapi belum di terima pemilik paket, resi JP7506189486 tolong di bantu cek min	2021-09-27 21.04	JNT
Sheptianoctav	biar belajar main PS pake PS5 #jntexpressid	2021-09-27 12.16	JNT

Pada proses scraping ini terdapat *flowchart* yang mana akan menjelaskan proses pengumpulan data atau *scraping tweet* diawali dengan memasukkan kata kunci atau *query* pencarian serta tanggal awal dan tanggal akhir, yang kemudian dilakukan pencarian *tweet* berdasarkan parameter menggunakan *library* Tweepy. Hasil pengumpulan data disimpan pada *data frame* lalu diekspor kedalam bentuk *file* Excel (.xlsx) sebelum kemudian data dipisahkan dan dimasukkan ke dalam *database*. *Flowchart* proses *scraping* dapat dilihat pada Gambar 1 berikut:



**Gambar 1.** *Flowchart* Scraping

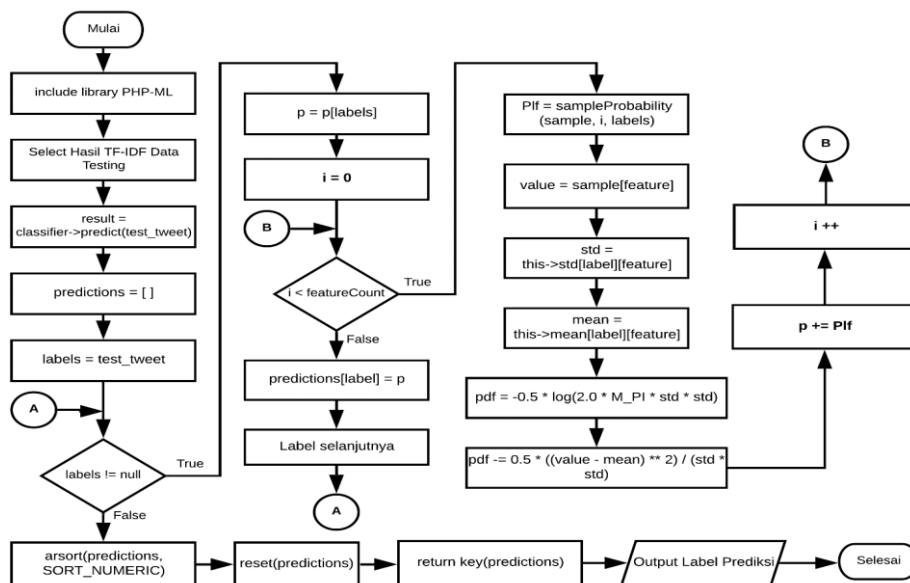
Dalam pengerjaan suatu sistem analisis sentimen twitter menggunakan metode Naïve Bayes, diperlukan gambaran umum yang dapat menunjukkan alur dari penelitian yang sedang dilakukan, agar dalam proses penelitian berjalan sesuai dengan rencana. Alur kerja ini dapat merepresentasikan proses tahapan awal hingga akhir sistem berjalan, gambaran dapat dilihat pada gambar 2. berikut:



**Gambar 2.** Tahap Keseluruhan Sistem

Pada gambar 2 terdapat beberapa tahapan dalam penelitian ini diantaranya teknik pengumpulan data menggunakan sistem scrapping data dari twitter untuk mendapatkan data set mentah. Selanjutnya Tweet yang sudah ada dalam bentuk file excel akan dilakukan proses sortir data menjadi data testing dan training, kemudian masing masing data akan diimport ke database melalui web untuk selanjutnya dilakukan preprocessing dan pelabelan kelas pada analisis sentimen tweet ini. Lalu pada tahap preprocessing akan dilakukan pemilihan kata yang tepat untuk selanjutnya di proses pada tahap selanjutnya, yaitu pelabelan kata sesuai analisis sentimen dari tweet yang digunakan sebagai acuannya. Lalu pada tahap klasifikasi Naive Bayes akan dilakukan pembobotan TF-IDF dan pemodelan. Selanjutnya dilakukan proses klasifikasi yang kemudian dilanjutkan dengan dilakukan pengujian dan perhitungan akurasi berdasarkan data yang ada untuk nantinya dapat mengetahui seberapa akurat dan presisi, serta hasil penerapan metode akan divisualisasikan ke dalam bentuk grafik dan presentase.

Pada *flowchart* ini, menjelaskan proses klasifikasi dengan algoritme *Naive Bayes Classifier*. Setelah model selesai dibuat dilanjutkan dengan proses klasifikasi dengan menggunakan hasil *weight* data uji, proses diawali dengan pembuatan *word vector* serta pembobotan TF-IDF data uji, sampai dengan menghitung prediksi label sentimen. *Flowchart* proses klasifikasi dapat dilihat pada Gambar 3. berikut:



**Gambar 3.** Flowchart Proses Klasifikasi



#### 4. IMPLEMENTASI

Agar aplikasi yang telah dikembangkan dapat berjalan dengan semestinya, dibutuhkanlah perangkat dengan spesifikasi tertentu, spesifikasi ini bertujuan untuk mendukung kinerja sistem agar dapat berjalan dengan baik dan memiliki hasil seperti yang diinginkan. Adapun spesifikasi perangkat yang digunakan dalam penelitian ini di antaranya:

Daftar perangkat keras yang mendukung aplikasi ini untuk berjalan dengan baik adalah sebagai berikut:

*Processor* : Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz  
*RAM* : 8 GB DDR 4  
*SSD* : 250 GB  
*VGA* : NVIDIA GeForce 930MX

Daftar perangkat lunak yang mendukung aplikasi ini untuk berjalan dengan baik adalah sebagai berikut:

*Sistem Operasi* : Windows 10 Enterprise 64-bit  
*DBMS* : MySQL Database  
*Web Server* : XAMPP (Apache v2.4.46, PHP v7.4.11)  
*Browser* : Google Chrome  
*IDE* : Visual Studio Code v1.53.0, & Jupyter Notebook

Algoritme adalah urutan atau alur tahapan proses yang dijabarkan dalam bentuk tulisan, algoritme juga merupakan representasi pengaplikasian dari *flowchart* yang telah dirancang sebelumnya dalam bentuk proses yang urut dan logis yang nantinya akan menjadi dasar dalam pembuatan kode program pada sistem ini.

Pada algoritme ini dijelaskan tentang proses keseluruhan sistem secara keseluruhan secara umum berdasarkan *flowchart* yang dibuat sebelumnya.

**Tabel 2.** Algoritme Keseluruhan Sistem

1. start
2. Lakukan proses Scraping
3. Baca dataset Tweet Kotor (Text)
4. Lakukan Pemisahan Data
5. Hasil Data Testing & Data Training
6. Insert Database
7. Lakukan Proses Preprocessing
8. Output Data Testing & Data Training Bersih
9. Lakukan Proses Pelabelan
10. Baca Data Training Bersih
11. Lakukan Pembobotan TF-IDF
12. Lakukan Proses Modeling
13. Output Model Training
14. Baca Data Testing Bersih
15. Lakukan Pembobotan TF-IDF
16. Lakukan Proses Klasifikasi
17. Output Prediksi Label Data Uji
18. end

Pada algoritme ini dijelaskan tentang tahapan pengumpulan data dengan cara *scraping* tweet.

**Tabel 3.** Algoritme Proses Scraping

1. start
2. input query, tanggal awal, dan tanggal akhir pencarian
3. lakukan proses Scraping
4. Output Dokumen Excel
5. end

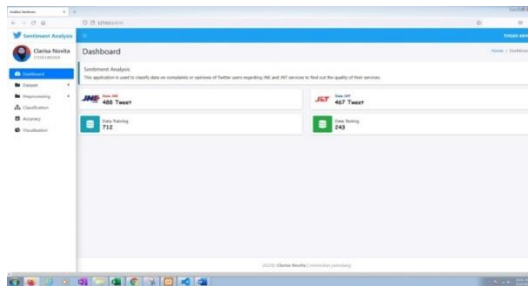
Pada algoritme ini dijelaskan tentang proses *preprocessing* yang dilakukan berdasarkan *flowchart* yang dibuat sebelumnya.

**Tabel 4.** Algoritme Proses Preprocessing

- |  |
|--|
| <ol style="list-style-type: none"> <li>1. start</li> <li>2. select dataset kotor (tweet)</li> <li>3. dictTweet = []</li> <li>4. Lakukan proses Casefolding (sts_lower)</li> <li>5. Lakukan proses Cleansing (sts_cleansing)</li> <li>6. Lakukan proses Normalization (tweet_string)</li> <li>7. Lakukan proses Stopword (outputstopword)</li> <li>8. Lakukan proses Stemming (outputstem)</li> <li>9. Output Tweet Bersih</li> </ol> |
| <ol style="list-style-type: none"> <li>10. Simpan outputstem ke dalam dictTweet</li> <li>11. Tweet Selanjutnya</li> <li>12. if(Tweet != null)</li> <li>13. kembali ke nomor 4</li> <li>14. else</li> <li>15. Simpan dictTweet ke dalam database</li> <li>16. endif</li> <li>17. end</li> </ol>   |

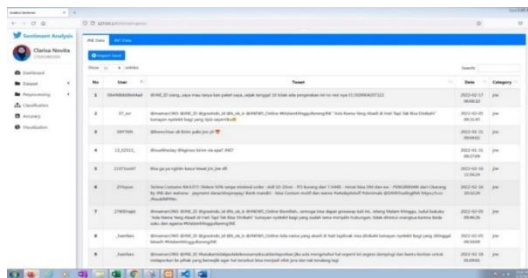
**a. Algoritme Proses Casefolding**

Pada algoritme ini dijelaskan tentang proses casefolding yang merupakan sub tahapan pada *processing*, yang dilakukan berdasarkan *flowchart* yang dibuat sebelumnya. Berikut ini merupakan tampilan layar dari aplikasi klasifikasi sentimen yang dibuat:



**Gambar 4.** Tampilan Layar *Dashboard*

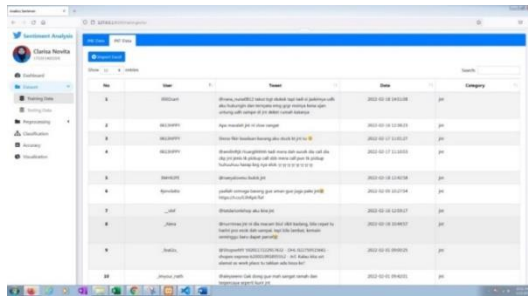
Pada Gambar 4 merupakan Tampilan layar dari Dashboard (Halaman Utama) yang berfungsi untuk melihat *dataset* yang sudah dimasukkan. Berikut tampilan aplikasi yang dibuat:



**Gambar 5.** Tampilan Layar *Dataset Training JNT*

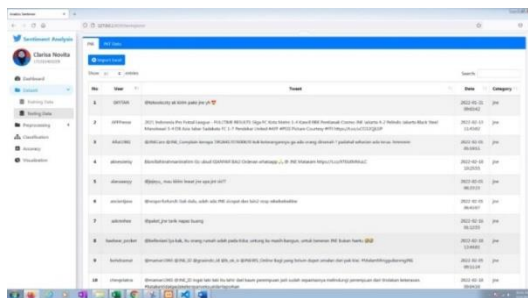
Pada Gambar 5 merupakan Tampilan layar dari *Training Data* yang berfungsi untuk menambahkan dan melihat *dataset*. Berikut tampilan aplikasi yang dibuat:





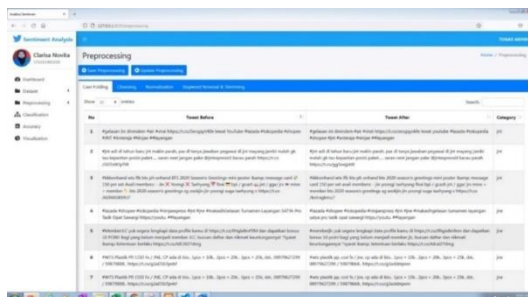
**Gambar 6.** Tampilan Layar Dataset Training JNE

Pada Gambar 6 merupakan Tampilan layar dari *Training Data* yang berfungsi untuk menambahkan dan melihat *dataset*. Berikut tampilan aplikasi yang dibuat:



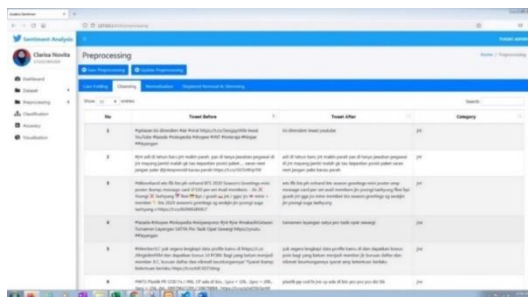
**Gambar 7.** Tampilan Layar Dataset Testing Data

Pada Gambar 7 merupakan Tampilan layar dari *Testing Data* yang berfungsi untuk menambahkan dan melihat *dataset*. Berikut tampilan aplikasi yang dibuat:



**Gambar 8.** Tampilan Layar Casefolding

Pada Gambar 8 merupakan Tampilan layar dari *Casefolding* yang berfungsi untuk menyimpan dan *update* hasil *preprocessing*. Berikut tampilan aplikasi yang dibuat:



**Gambar 9.** Tampilan Layar Cleaning

Pada Gambar 8 merupakan Tampilan layar dari *Cleansing* yang berfungsi untuk menyimpan dan *update* hasil *preprocessing*. Berikut tampilan aplikasi yang dibuat:

## 5. KESIMPULAN

Berdasarkan dari penelitian yang telah penulis lakukan yang melalui tahapan pembuatan system serta pengujian langsung dari aplikasi, dengan menggunakan data - data yang telah ada dan menggunakan metode algoritma yang ada, maka penulis menyimpulkan beberapa hal sebagai berikut:

1. Hasil dari klasifikasi tweet sentiment positif, negative dan netral dari kalimat yang telah ada untuk masing masing jasa pengiriman telah diperoleh bahwa JNE mendapat 0 respon positif, 18 respon negative dan 56 respon netral. Sedangkan JNT mendapat 0 respon positif, 2 respon netral dan 167 respon negative
2. Hasil dari pengujian yang telah dilakukan oleh peneliti bahwa beberapa pengguna lebih senang menggunakan jasa kirim JNE daripada JNT, karena JNT lebih banyak mendapatkan respon negative dari pengguna layanan jasa kirim

Hasil dari pengujian yang menggunakan metode naïve bayes dengan rumus Densitas Gauss pada data kontinu menghasilkan akurasi 58.02 %, presisi 43.33%, recall 41.67%.

## REFERENCES

- Joko Suntoro. (2019). Data Mining : Algoritma dan Implementasi dengan Pemrograman PHP.
- Afrizal, S., Irmanda, H. N. and Falih, N. (2019) 'Implementasi Metode Naïve Bayes untuk Analisis Sentimen Warga Jakarta Terhadap Kehadiran Mass Rapid Transit', 4221, pp. 157–168.
- Agusta, L. (2009) 'Perbandingan Algoritma Stemming Porter Dengan Algoritma Nazief & Adriani Untuk Stemming Dokumen Teks Bahasa Indonesia', *Konferensi Nasional Sistem dan Informatika 2009*, (KNS&I09-036), pp. 196–201.
- Andika, L. A., Azizah, P. A. N. and Respatiwan, R. (2019) 'Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier', *Indonesian Journal of Applied Statistics*, p. 34. doi: 10.13057/ijas.v2i1.29998.
- APJII (2020) 'Laporan survei internet apjii 2019 – 2020', *Asosiasi Penyelenggara Jasa Internet Indonesia*, 2020, pp. 1–146.
- Attabi, A. W., Muflikhah, L. and Fauzi, M. A. (2018) 'Penerapan Analisis Sentimen untuk Menilai Suatu Produk pada Twitter Berbahasa Indonesia dengan Metode Naïve Bayes Classifier dan Information Gain', *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(11), pp. 4548–4554.
- Fauziyyah, A. K. (2020) 'Analisis Sentimen Pandemi Covid19 Pada Streaming Twitter Dengan Text Mining Python', *Jurnal Ilmiah SINUS*, 18(2), p. 31. doi: 10.30646/sinus.v18i2.491.
- Hadna, M. S., Santosa, P. I. and Winarno, W. W. (2016) 'Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter', *Seminar Nasional Teknologi Informasi dan Komunikasi*, 2016(Sentika), pp. 57–64.
- Mujilawati, S. (2016) 'Pre-Processing Text Mining Pada Data Twitter', *Seminar Nasional Teknologi Informasi dan Komunikasi*, 2016(Sentika), pp. 2089–9815.
- Mustofa Hidayat, A. and Syafrullah, M. (2017) 'Algoritma Naïve Bayes Dalam Analisis Sentimen Untuk Klasifikasi Pada Layanan Internet PT.XYZ', *Jurnal TELEMATIKA MKOM*, 9(2), pp. 91–95.
- Oktasari, L., Chrisnanto, Y. H. and Yuniarti, R. (2016) 'Text Mining Dalam Analisis Sentimen Asuransi Menggunakan Metode Naive Bayes Classifier', *Prosiding SNST*, 7, pp. 37–42.
- Putra, R. A. (2019) 'Penerapan Naïve Bayes Classifier dengan Gaussian Function Untuk Menentukan Kelompok UKT', *Jurnal Ilmiah Informatika Global*, 9(2), pp. 112–117. doi: 10.36982/jig.v9i2.583.
- Rahutomo, F. and Ririd, A. R. T. H. (2019) 'Evaluasi Daftar Stopword Bahasa Indonesia', *Jurnal Teknologi Informasi dan Ilmu Komputer*, 6(1), p. 41. doi: 10.25126/jtiik.2019611226.
- Rustiana, D. and Rahayu, N. (2017) 'Analisis sentimen pasar otomotif mobil', *Jurnal SIMETRIS*, 8(1), pp. 113–120.



- Saleh, A. (2015) 'Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga', *Creative Information Technology Journal*, 2(3), pp. 207–217.
- Santoso, E. B. and Nugroho, A. (2019) 'Analisis Sentimen Calon Presiden Indonesia 2019 Berdasarkan Komentar Publik Di Facebook', *Eksplora Informatika*, 9(1), pp. 60–69. doi: 10.30864/eksplora.v9i1.254.
- Sudiantoro, A. V. and Zuliarso, E. (2018) 'Analisis Sentimen Twitter Menggunakan Text Mining Dengan Algoritma NAÏVE BAYES CLASSIFIER', *Prosiding SINTAK 2018*, 10(2), pp. 398–401.
- Tala, F. Z. (2003) 'A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia', *M.Sc. Thesis, Appendix D*, pp. 39–46.
- Wahid, D. H. and SN, A. (2016) 'Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity', *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 10(2), p. 207. doi: 10.22146/ijccs.16625.
- Wahyudi, D., Susyanto, T. and Nugroho, D. (2017) 'Implementasi Dan Analisis Algoritma Stemming Nazief & Adriani Dan Porter Pada Dokumen Berbahasa Indonesia', *Jurnal Ilmiah SINUS*, 15(2), pp. 49–56. doi: 10.30646/sinus.v15i2.305.
- Zulfa, I. and Winarko, E. (2017) 'Sentimen Analisis Tweet Berbahasa Indonesia Dengan Deep Belief Network', *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 11(2), p. 187. doi: 10.22146/ijccs.24716.