

ANALISIS SENTIMEN TERHADAP PELUNCURAN MACBOOK MODEL TERBARU PADA TWITTER DENGAN METODE NAIVE BAYES

Fiqih Alfiansyah Zahari ^{a,1,*}, Abdul Halim Anshor ^{b,2}, Sugeng Budi Rahardjo ^{c,3}

^{a, b, c} Universitas Pelita Bangsa, Jl. Inspeksi Kalimalang No.9, Cibatu, Cikarang Selatan, Bekasi 17530, Indonesia

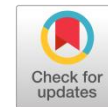
¹ fiqihalfiansyah6@gmail.com *; ² abduhalimanshor@pelitabangsa.ac.id; ³ sugeng@pelitabangsa.ac.id

* Penulis Korespondensi

Diterima : 12 Agustus 2023 | Direvisi : 15 Agustus 2023 | Diterbitkan : 18 Agustus 2023

ABSTRAK

Sudah banyak berbagai jenis laptop beredar dipasaran, yang saat ini terbagi menjadi 3 kelas, yaitu low-end, mid-end, dan high-end yang memiliki kelebihan masing-masing. Twitter merupakan salah satu media sosial yang populer di kalangan pengguna internet, hal ini dikarenakan twitter mudah digunakan, bebas dalam mengeluarkan pendapat atau opini dan membahas berbagai masalah yang ada. Masalah yang banyak dibahas adalah peluncuran Macbook model terbaru salah satunya. Permasalahan ini mendapat perhatian dari masyarakat luas sehingga opini tersebut dapat diolah dengan cara yaitu analisis sentimen dan diperlukan sebuah metode yang dapat secara otomatis melakukan klasifikasi opini ke dalam kategori positif dan negatif melalui proses analisis sentimen. Algoritma ini bekerja berdasarkan prinsip probabilitas bersyarat, seperti yang diberikan oleh Teorema Bayes. Teorema Bayes adalah metode untuk menemukan probabilitas ketika kita mengetahui probabilitas tertentu lainnya. Naive Bayes merupakan teknik machine learning yang populer untuk pengklasifikasian. Klasifikasi dapat memberikan kemudahan bagi pengguna untuk melihat opini positif dan negatif. Berdasarkan pengujian yang telah dilakukan, hasil klasifikasi terbaik diperoleh dengan nilai accuracy 84.53%, precision 99.79%, dan recall 75.70%.

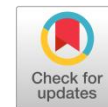


KATA KUNCI

Macbook
Klasifikasi
Opini
Data Mining
Naive Bayes

ABSTRACT

There are many different types of laptops on the market, which are currently divided into 3 classes, namely low-end, mid-end and high-end, each of which has its advantages. Twitter is one of the most popular social media among internet users, this is because Twitter is easy to use, free to express opinions or opinions and discuss various existing problems. The problem that has been widely discussed is the launch of the latest Macbook model, one of which. This problem has received attention from the wider community so that these opinions can be processed by means of sentiment analysis and we need a method that can automatically classify opinions into positive and negative categories through the sentiment analysis process. This algorithm works on the conditional probability principle, as given by Bayes Theorem. Bayes theorem is a method for finding probabilities when we know certain other probabilities. Naive Bayes is a popular machine learning technique for classification. Classification can make it easy for users to see positive and negative opinions. Based on the tests that have been done, the best classification results are obtained with accuracy values of 84.53%, precision of 99.79%, and recall of 75.70%.



KEYWORD

Macbook
Classification
Opinion
Data Mining
Naive Bayes



This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license

1. Pendahuluan

Di era teknologi saat ini, salah satu diantaranya perkembangan internet. Perkembangan internet yang cepat telah mengubah banyak aspek dalam proses komunikasi data komputer[1]. Dengan hadirnya internet sebagai fasilitas yang sangat dibutuhkan membuat semakin banyaknya platform-platform yang berkembang dan bisa kita jumpai termasuk platform media sosial[2]. Sudah banyak berbagai jenis laptop beredar dipasaran, yang saat ini terbagi menjadi 3 kelas, yaitu *low-end*, *mid-end*, dan *high-end* yang memiliki kelebihan masing-masing, seperti laptop *low-end* yang memiliki harga yang terjangkau tetapi spesifikasi

yang kurang mempunyai seperti *Read Only Memory* (ROM) yang masih menggunakan *Hard Disk Drive* (HDD) kurang lebih hanya 250GB saja, ada juga laptop *mid-end* yang biasanya sudah bisa dipakai untuk *video editing* tetapi memakan waktu yang cukup lama untuk merender *video*, dan juga ada laptop *high-end* yang memiliki spesifikasi tinggi sebagai contoh ada Macbook Pro M1 yang sudah mendukung hingga resolusi 1680 x 1050 dengan Chip Apple M1 dengan daya tahan baterai hingga mencapai 20 jam dan juga sudah menggunakan *Solid State Drive* (SSD) sebesar 256GB dengan spesifikasi seperti ini merender video jauh lebih cepat dibanding laptop *mid-end*. Sudah memiliki spesifikasi yang tinggi, Macbook juga mempunyai desain yang elegan dan dibuat dengan material premium terbaik, didalam spesifikasi yang tinggi serta desain yang elegan macbook memiliki harga yang cukup tinggi kisaran Rp. 12.000.000 – Rp. 16.000.000 dengan harga segitu target pasar Macbook dikhususkan untuk masyarakat *high class* yang membuat masyarakat *mid class* kebawah memikirkan ulang untuk membelinya karena harga yang sangat tinggi.

Analisis sentimen merupakan analisa terhadap suatu peristiwa dari pendapat yang didasarkan pada sikap seseorang tentang suatu objek[3]. Analisis sentimen biasanya dilakukan untuk mengumpulkan dan mengetahui opini masyarakat dalam postingan twitter. Analisis sentimen dibutuhkan dengan tujuan untuk mengetahui opini publik terhadap suatu objek[4]. Oleh karena itu dibutuhkan suatu analisis terhadap opini-opini tersebut dalam penelitian ini agar bisa dijadikan tolak ukur baik atau tidaknya Macbook[5]. Opini-opini tersebut bisa berupa opini negatif atau positif tergantung dari pandangan publik terhadap objek tersebut. Penggunaan metode Naive Bayes ini, dianggap tepat karena metode ini mudah diterapkan untuk mencari *trending event*. Pengklasifikasian yang saling bebas membuat peluang tiap parameter menjadi lebih mudah. Naive Bayes merupakan sebuah metode klasifikasi yang berakar pada Teorema Bayes. Ciri utama dari Naive Bayes adalah asumsi yang sangat kuat akan independensi dari masing-masing kondisi atau kejadian[6].

2. Tinjauan Pustaka

Adapun beberapa referensi yang berkaitan dengan analisis sentimen menggunakan algoritma naive bayes sebagai berikut:

(Nila Hardi, Yuris Alkahfi, Popon Handayani, Windu Gata, dan Muhammad Rifqi Firdaus, 2020) Analisis Sentimen *Physical Distancing* pada Twitter Menggunakan *Text Mining* dengan Algoritma Naive Bayes Classifier. Pada perhitungan analisis sentimen terhadap *physical distancing* di tengah pandemi covid-19 menggunakan metode NBC memperoleh hasil akurasi sebesar 50,26%. Tujuan dari penelitian ini, agar dapat mengkategorikan opini negatif atau positif, dari pembahasan *physical distancing*. Nantinya informasi terkait kebijakan *physical distancing* bisa sampai tepat informasinya kepada masyarakat[7].

(Imam Fahrur Rozi, Annisa Taufika Firdausi, dan Khalimatul Islamiyah, 2020) Analisis Sentimen pada Twitter Mengenai Pasca Bencana Menggunakan Metode Naive Bayes dengan Fitur N-Gram. Pada penelitian ini, dilakukan analisis sentimen terhadap tweet dengan mengklasifikasikan kedalam kategori positif atau negatif menggunakan algoritma Naive Bayes Classifier. Berdasarkan hasil klasifikasi tersebut dapat diprioritaskan bantuan dengan mengkategorikan setiap hasil klasifikasi. Dari hasil klasifikasi tersebut positif diartikan bantuan berdasarkan kategori tersebut telah terpenuhi sedangkan negatif berarti bantuan tersebut kurang atau dibutuhkan. Penelitian ini juga akan menggunakan fitur N-Gram yaitu unigram dan bigram. Pada pengujian penelitian ini dilakukan empat kali pengujian. Setiap pengujian, persentase data *training* berbeda karena jumlah data *training* juga berpengaruh terhadap peningkatan nilai akurasi. Hasil dari pengujian tersebut diperoleh nilai akurasi untuk unigram sebesar 76.67%, 84.44%, 90.00% dan 93.33%. Nilai akurasi untuk bigram sebesar 64.17%, 68.89 %, 75.00%, 86.67%. Dari empat pengujian didapatkan hasil akurasi tertinggi pada unigram yaitu sebesar 93.33% dan bigram sebesar 86.67%. Jadi untuk nilai akurasi unigram lebih tinggi daripada bigram[8].

(Lingga Aji Andika, Pratiwi Amalia Nur Azizah, Respatiwan, 2019) Analisis Sentimen Masyarakat terhadap Hasil *Quick Count* Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier. Kesimpulan dari penelitian yang telah dilakukan adalah bahwa pengujian memperoleh tingkat akurasi terbaik sebesar 82,90% dengan $\alpha = 0,05$ yang berarti bahwa data dapat diklasifikasikan kedalam positif ataupun negatif secara benar sebanyak 82,90% dengan tingkat signifikansi 0,05. Kemudian didapatkan data sebanyak 34,5% (471) tweet positif dan 65,5% (895) tweet negatif terhadap hasil quick count[9].

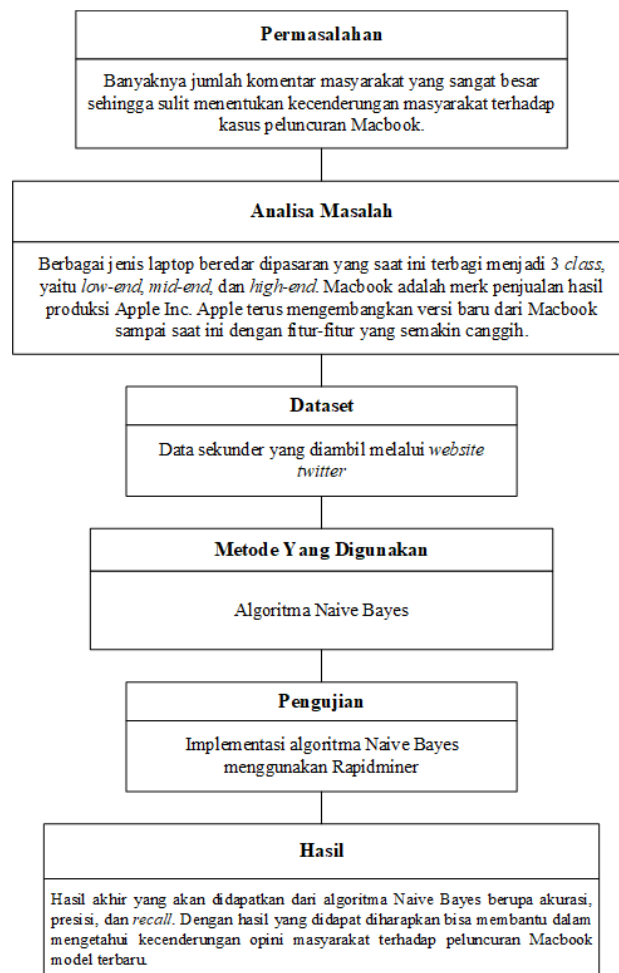
(Winda Yulita, Eko Dwi Nugroho, Muhammad Habib Algifari, 2021) Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naive Bayes Classifier. Berdasarkan hasil analisis yang dilakukan pendapat masyarakat Indonesia khususnya pengguna media sosial Twitter, rata-rata memberikan respon positif terkait adanya kebijakan vaksinasi Covid-19 di Indonesia dibuktikan dengan persentase respon positif sebanyak 60.3% (2278 respon). Penggunaan algoritma Naive Bayes Classifier untuk melakukan analisis sentimen ini sudah sangat baik ditunjukkan dengan hasil akurasi yang tinggi sebesar 93%[10].

(Fransiska Vina Sari, Arief Wibowo, 2019) Analisis Sentimen Pelanggan Toko Online Jd.id Menggunakan Metode Naive Bayes Classifier Berbasis Konversi Ikon Emosi. Kesimpulan dari penelitian yang sudah dilakukan NBC tanpa pembobotan tf-idf dan konversi ikon emosi (*convert emoticon*) memiliki nilai akurasi 96,44% sedangkan NBC dengan pembobotan tf-idf dan konversi ikon emosi (*convert emoticon*) memiliki nilai akurasi 98%, dengan demikian terjadi peningkatan akurasi sebesar 1,56%[11].

3. Metodologi Penelitian

3.1. Kerangka Pemikiran

Kerangka pemikiran ini dibuat untuk mempermudah memahami arah penelitian dari proses-proses hingga hasil yang akan diperoleh, kerangka berfikir ini dapat dilihat pada gambar berikut:



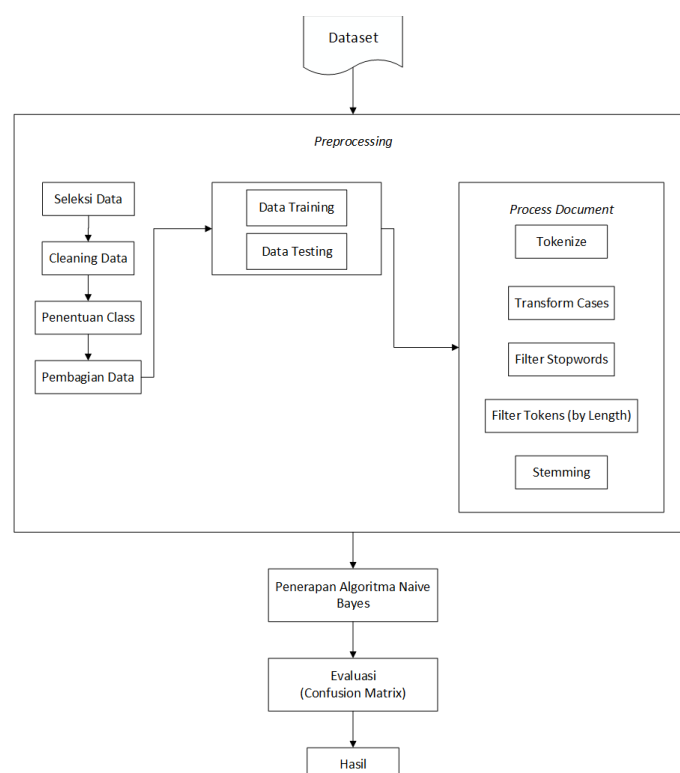
Gambar 1. Kerangka Pemikiran

Menjelaskan permasalahan yang ada sebagai latar belakang penelitian dan tujuan yang akan dicapai, menganalisa masalah dan membuat ruang lingkup masalah agar lebih sederhana, data yang digunakan merupakan data sekunder yang diambil melalui pihak yang telah mengumpulkan data tersebut sebelumnya, setelah itu melakukan pengujian terhadap dataset menggunakan algoritma Naive Bayes.

Objek penelitian merupakan isu atau permasalahan yang dibahas dan diteliti dalam penelitian. Peluncuran Macbook model terbaru mendapat perhatian dari masyarakat luas tentang pro maupun kontra. Maka dari itu objek penelitian ini adalah respon masyarakat terhadap peluncuran Macbook tersebut. Subjek dalam penelitian ini adalah Apple Inc. selaku produksi yang akan bertanggung jawab terkait peluncuran Macbook model terbaru. Data yang diambil berasal dari media sosial twitter dengan *keyword* “macbook” dan data yang diambil berupa komentar masyarakat pada twitter. Data yang digunakan merupakan data yang diperoleh dari twitter, baik berupa tweet atau komentar yang diambil melalui proses *crawling*. Data yang sudah terkumpul nantinya akan dibagi menjadi dua bagian yaitu data latih dan data uji.

3.2. Metode Penelitian

Penelitian ini akan melalui beberapa proses, urutan proses penelitian ini akan diuraikan pada metode penelitian ini:



Gambar 2. Metode Penelitian

Dataset didapat dari hasil *crawling* data menggunakan *tools* RapidMiner. Dataset merupakan data teks berbahasa Indonesia yang diambil melalui media sosial twitter dengan *keyword* “macbook” dan data diambil sebanyak 3000 data. Setelah itu dilakukan tahap *preprocessing* untuk seleksi data, membersihkan simbol-simbol didalam data yang tidak diperlukan, penentuan *class*, dan juga pembagian data *training* dan data *testing* menggunakan Split Data dengan ratio 90:10 antara data *training* dan data *testing*. Selanjutnya masuk ke tahap *Process Document from Data* untuk melakukan normalisasi data. Setelah itu penerapan algoritma Naive Bayes terhadap dataset untuk mendapatkan hasil dari *Confusion Matrix*.

4. Hasil dan Pembahasan

4.1. Hasil

a. *Crawling Data*

Pengumpulan data dilakukan dengan cara *crawling* data menggunakan perangkat lunak RapidMiner. Jumlah data *tweet* yang diperoleh dari proses tersebut sebanyak 3000, data yang dikumpulkan merupakan data teks berbahasa Indonesia yang diambil dari akun media sosial twitter dengan *keyword* "macbook". Berikut adalah hasil dari *crawling* data:

Row No.	Created-At	From-User	From-User-Id	To-User	To-User-Id	Language	Source	Text	Geo-Location	Ge
1	Oct 31, 2022 ...	????	632874431	?	-1	in	<a href="http://...	Sekarang cult...	?	?
2	Oct 31, 2022 ...	Whiskey	3196849650	roangdungle	9477540000...	in	<a href="http://...	Menurut gue ...	?	?
3	Oct 31, 2022 ...	vin	1412810000...	?	-1	in	<a href="http://...	macbook gua...	?	?
4	Oct 31, 2022 ...	Gadget Diva ID	272841361	?	-1	in	<a href="http://...	Hingga akhir ...	?	?
5	Oct 31, 2022 ...	anak rumahan	1439950000...	?	-1	in	<a href="https://...	LAPTOP SLE...	?	?
6	Oct 31, 2022 ...	?	3065993046	?	-1	in	<a href="http://...	gimana sih c...	?	?
7	Oct 31, 2022 ...	????? ??-??...	1437970000...	EtmintR	1464610000...	in	<a href="https://...	itu dia minta ...	?	?
8	Oct 31, 2022 ...	cotton candy	1522270000...	EtmintR	1464610000...	in	<a href="http://...	itu dia mau m...	?	?
9	Oct 31, 2022 ...	khushi singh	1564530000...	?	-1	in	<a href="https://...	iPad Repair i...	?	?
10	Oct 31, 2022 ...	khushi singh	1564530000...	?	-1	in	<a href="https://...	iPhone Repai...	?	?
11	Oct 31, 2022 ...	???	96539382	cucul	38735811	in	<a href="http://...	Di saya gak a...	?	?
12	Oct 31, 2022 ...	khushi singh	1564530000...	?	-1	in	<a href="https://...	iMac Repair i...	?	?
13	Oct 31, 2022 ...	khushi singh	1564530000...	?	-1	in	<a href="https://...	Macbook Rep...	?	?
14	Oct 31, 2022 ...	COLLE CEK...	1046080000...	?	-1	in	<a href="https://...	hai, kalian ad...	?	?

Gambar 3. Hasil Proses *Crawling*

b. *Seleksi Data*

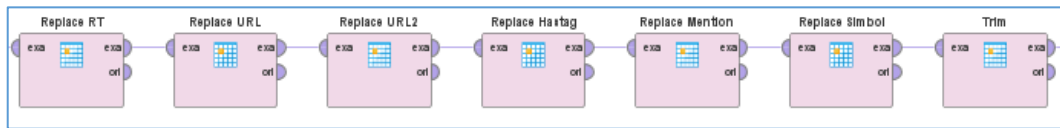
Pada tahap ini bertujuan untuk menyeleksi data yang akan digunakan, dikarenakan data tweet yang didapat dari proses *crawling* masih memiliki data-data yang tidak diperlukan, dan pada penelitian ini data yang diambil hanya membutuhkan komentar tweet atau data yang ada pada kolom text.

Text
terus sebulan setelah sempro dikasih kabar sama bokap kalo bos nya ngasih dia macbook sama iphone, kekk ya Allah bokap gua berdoa nya sekenceng apa ya sampe
[u]f halo boleh minta tolong ga, ada temen yang MacBook nya ketinggalan di ruang kelas 3404 di gedung 3. Kalau ada yang liat boleh dm ya
buat yang ingin tampil kekinian, kini anda bisa sewa iphone, sewa ipad, sewa apple tv, sewa imac, sewa macbook harian/mingguan/bulanan. 085271622233
@ndagels https://t.co/BnxU0fMbc cek link macbook bekas under 10 jt disini ya gais. barangkali ada yang lagi nyari
@collegemenfess Aku pakai epson sih nder, awet ?? kalau pakai adapter, tergantung tipe macbook kamu. Aku macbook air 2017 gaperlu adapter buat printer
@01shkar macbookkishi peya
Macbook nalang talaga. ?
pagi, siang, sore ngantuk mulu bawaannya. pas malem tiba-tiba pengen iPad, Macbook, iMac.
@UGM_FESS Tanya dong, biar bisa ngedit di aplikasi microsoft (word, dsb) di macbook gimana ya? Udah masuk akun microsoft 365 ugm juga. Apa harus hubungin a
sewa macbook pro di jakarta? klik https://t.co/gXAs2cOSJt +6285271622233 email https://t.co/t2qfElqWdn
@repeurep @tanyarifes Nah ini juga anjir, screenshot iPhone sama MacBook
sewa macbook untuk editing video di jakarta? klik https://t.co/gXAs2cOSJt +6285271622233 email https://t.co/t2qfElqWdn
@infomojokerto_ tolong jika ada yg menemukan seseorang menjual unit macbook air 2019 beritahu saya. Adik saya kehilangan laptopnya di bis yang kata knek bis ny
tak hairan pon dulu time study ada yang tiap minggu tukar macbook baru, katanya storage penuh wow sangat apa tk boleh reformat ke? https://t.co/2WYsXhrBBw
RT @Kantenn3: Anak sulit memastikan BLE Tag with mac address berasal dr orang yg di aksesin, karena baharuna peralatan yg support BLE iuna hr

Gambar 4. Hasil Seleksi Data

c. Cleaning Data

Proses *cleaning* adalah tahapan untuk menghapus kata, simbol, dan karakter yang tidak diperlukan pada penelitian ini.



Gambar 5. Subprocess Cleaning Data

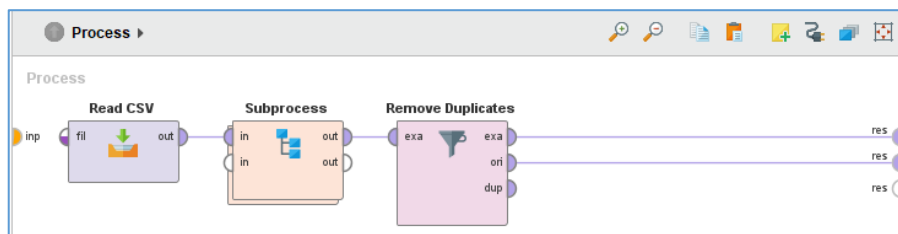
Operator *Subprocess* digunakan untuk menampung setiap proses menjadi satu agar terlihat lebih rapih. Terdapat Operator Replace yang dibagi menjadi 6 bagian, Operator Replace RT digunakan untuk menghapus kata retweet, Operator Replace URL 1 dan 2 digunakan untuk menghapus (Uniform Resource Locator), Operator Replace Hastag digunakan untuk menghapus simbol (#), Operator Replace Mention digunakan untuk menghapus simbol (@), Operator Replace Simbol digunakan untuk menghapus simbol-simbol yang tidak diperlukan, dan Operator Trim digunakan untuk menghapus spasi yang tidak diperlukan.

Row No.	Text
1	Engga fast charging si setau aku temenku suka pinjem charger macbook aku kalo ga bawa charger
2	Kondisi bekas rasa baru
3	ngga anjir wkwk
4	oo saya cinta pada macbook
5	Ramai dah yang apply lulus ansuran iPhone series ni Student pun ramai yang lulus ipad MacBook
6	terlalu nanggung nder
7	Macbook Pro inch MD second cukup sih
8	lebih recomend macbook sih jujurr
9	Weh macbook aku bagus
10	rek macbook pro buat kuliah sama editing ringan masih worth to buy gak ya
11	Kemarin macbook sekarang kok windows tan
12	bismillah kebeli macbook pro
13	itu macbook pro masi worth gak rek buat kuliah dan editing ringan
14	Udah pake MacBook Pro M masih ga bagus

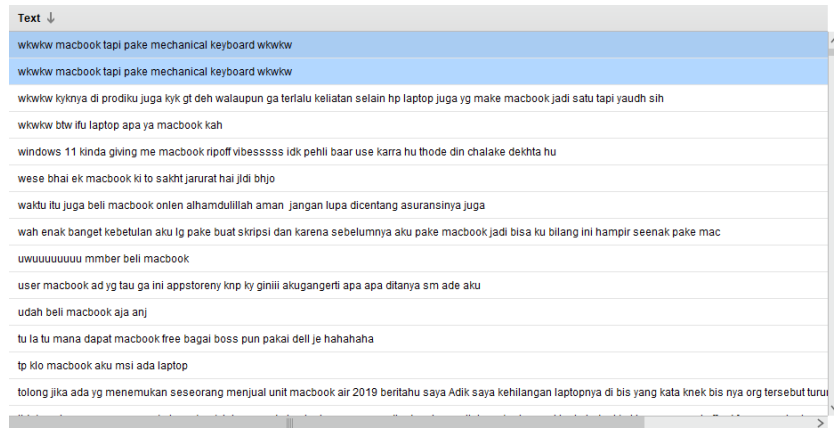
Gambar 6. Hasil Cleaning Data

d. Remove Duplicate

Proses ini bertujuan untuk memfilter data tweet yang sama, hal ini disebabkan karena twitter memiliki fitur retweet yang menyebabkan banyaknya tweet yang sama di retweet oleh pengguna lain dan menimbulkan teks berulang dengan topik yang sama maka dari itu dibutuhkan Operator *Remove Duplicates* untuk menghilangkan data yang *double*. Berikut adalah contoh tweet yang sama:



Gambar 7. Process Remove Duplicate



Gambar 8. Contoh tweet yang sama

e. Penentuan *Class Attribute*

Data tweet yang telah diproses dengan tahapan-tahapan sebelumnya akan dilakukan penentuan kelas atributnya, penentuan kelas atribut diberikan secara manual apakah itu termasuk sentimen positif atau sentimen negatif, berikut ini adalah proses penentuan kelas atribut dan dokumen yang telah ditentukan sentimennya:

Tabel 1. Proses penentuan *Class Attribute*

Tweet	Sentimen
Shock banget ini gatau fitur dari iOS atau macOS tapi connect wifi cukup dari macbook terus iphone nya otomatis connect ke wifi itu juga tanpa perlu pencet connect input password lagi	Positif
Pengen macbook tapi ya buat apa wkwk	Negatif

f. Pembagian Data

Jumlah data tweet yang diperoleh dari proses *crawling* data sebanyak 3000 data. Setelah dilakukan *preprocessing* pada tahapan sebelumnya diperoleh 1015 data yang sudah menjadi data bersih dan juga diberi label sentimen secara manual. Pembagian data bertujuan untuk memperoleh data *training* dan data *testing*, data tersebut nantinya akan dibagi menggunakan operator Split Data dengan ratio 90% sebagai data latih (*training*) dan 10% sebagai data uji (*testing*) atau masing-masing 914 data sebagai data latih (*training*) dan 101 data sebagai data uji (*testing*). Berikut dataset yang akan digunakan untuk pengujian:

Row No.	Sentimen	Text
1	Positif	Engga fast charging si setau aku temenku suka pinjem charger macbook aku kalo ga bawa charger
2	Positif	Kondisi bekas rasa baru
3	Negatif	ngga anjir wkwk
4	Positif	oo saya cinta pada macbook
5	Positif	Ramai dah yang apply lulus ansuran iPhone series ni Student pun ramai yang lulus ipad MacBook
6	Negatif	terlalu nanggung nder
7	Positif	Macbook Pro inch MD second cukup sih
8	Positif	lebih recomend macbook sih jujur
9	Positif	Weh macbook aku bagus
10	Positif	rek macbook pro buat kuliah sama editing ringan masih worth to buy gak ya
11	Negatif	Kemarin macbook sekarang kok windows tan
12	Positif	bismillah kebeli macbook pro
13	Positif	itu macbook pro masi worth gak rek buat kuliah dan editing ringan
14	Positif	Udah pake MacBook Pro M masih ga bagus

ExampleSet (1,015 examples, 1 special attribute, 1 regular attribute)

Gambar 9. Keseluruhan Dataset

g. Dataset

Jumlah dataset yang diperoleh dari proses *crawling* data sebanyak 3000 data. Setelah dilakukan proses seperti tahapan-tahapan diatas diperoleh 1015 data yang mana data itu terdiri dari data sentimen positif sebanyak 642 data dan sentimen negatif sebanyak 373 data, untuk lebih jelasnya dapat dilihat pada tabel dataset berikut:

Tabel 2. Dataset

Keyword	Data Crawling	Data Akhir	Positif	Negatif
macbook	3000	1015	642	373

h. Penggabungan *Attribute Data Training* dan *Data Testing*

Pada tahap ini akan dilakukan penggabungan antara atribut data *training* atau data latih yang akan di uji dengan data *testing* atau data uji yang sebelumnya sudah kita tentukan sentimennya secara manual. Adapun tahapan nya sebagai berikut:

1. *Union*

Ketika dataset antara data *training* dan data *testing* tidak lengkap maka Union akan menggabungkan atribut keduanya dan mengisinya sebagai *Missing Values* atau kolom yang tidak ada nilainya.

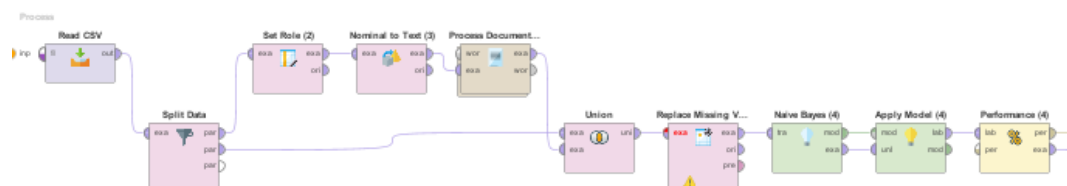
2. *Replace Missing Values*

Operator ini digunakan untuk mengganti nilai yang tidak ada pada data ini menjadi nilai nol.

4.2. Pembahasan

1. Pengujian Dataset Dengan Algoritma Naive Bayes

Dalam melakukan pengujian dataset, data dibagi menjadi dua yaitu data training dan data testing. Proses pengujian ini merangkai beberapa operator proses yang sesuai yaitu dengan Read CSV, Split Data, Set Role, Nominal to Text, Process Document from Data, Union, Replace Missing Values, Naive Bayes, Apply Model, Performance Classification. Proses pengujian dapat dilihat pada gambar berikut:



Gambar 10. Process Pengujian Algoritma Naive Bayes

2. Hasil Pengujian Dengan Algoritma Naive Bayes

Performa atau akurasi merupakan tingkat kedekatan. Pada hasil pengujian ini akan dihitung tingkat *accuracy*, *precision*, dan *recall* menggunakan RapidMiner untuk mengetahui seberapa akurat hasil pengujian pada *dataset* tersebut. Pada penelitian ini dilakukan satu kali pengujian untuk menghitung tingkat *accuracy*, *precision*, dan *recall* dengan *ratio* data *training* 90% dan data *testing* 10%, yaitu sebagai berikut:

accuracy: 84.53%

	true Negatif	true Positif	class precision
pred. Negatif	372	156	70.45%
pred. Positif	1	486	99.79%
class recall	99.73%	75.70%	

Gambar 11. Hasil Confusion Matrix

3. Accuracy

Adapun hasil *accuracy* diperoleh dengan cara menjumlah data *true positive* (TP) ditambah dengan data *true negative* (TN) dibagi dengan jumlah data *true positive* (TP) ditambah dengan data *true negative* (TN) ditambah dengan data *false positive* (FP) ditambah dengan data *false negative* (FN) Berikut cara perhitungan manual dari *accuracy*:

$$\begin{aligned}
 &= \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \\
 &= \frac{486+372}{486+372+1+156} \times 100\% \\
 &= \frac{858}{1015} \times 100\% \\
 &= 84.53\%
 \end{aligned}$$

4. Precision

Adapun hasil *precision* diperoleh dengan cara data *true positive* (TP) dibagi dengan jumlah data *true positive* (TP) ditambah dengan data *false positive* (FP). Berikut cara perhitungan manual dari *precision*:

$$\begin{aligned}
 &= \frac{TP}{TP+FP} \times 100\% \\
 &= \frac{486}{486+1} \times 100\% \\
 &= \frac{486}{487} \times 100\% \\
 &= 99.79\%
 \end{aligned}$$

5. Recall

Adapun hasil *recall* diperoleh dengan cara data *true positive* (TP) dibagi dengan jumlah data *true positive* (TP) ditambah dengan data *false negative* (FN). Berikut cara perhitungan manual dari *recall*:

$$\begin{aligned}
 &= \frac{TP}{TP+FN} \times 100\% \\
 &= \frac{486}{486+156} \times 100\% \\
 &= \frac{486}{642} \times 100\% \\
 &= 75.70\%
 \end{aligned}$$

5. Penutup

5.1. Kesimpulan

Berdasarkan hasil penelitian yang sudah dilakukan sebelumnya maka dapat diambil kesimpulan bahwa penerapan algoritma Naive Bayes dalam melakukan analisis sentimen dapat diterapkan dengan baik dan memberikan hasil yang sangat baik sehingga dapat membantu dalam mengetahui kecenderungan opini masyarakat terhadap peluncuran macbook model terbaru.

Dari proses pengujian yang dilakukan dengan bobot data *training* 90% dan data *testing* 10% mendapatkan hasil *accuracy* sebanyak 84.53%, *precision* sebanyak 99.79%, dan *recall* sebanyak 75.70%. Diperoleh jumlah data tweet dengan polaritas sentimen positif lebih banyak dibandingkan dengan data tweet dengan polaritas sentimen negatif. Hasil sentimen pada tweet macbook memperoleh sentimen positif 63% dan sentimen negatif 37% sehingga dapat disimpulkan bahwa kecenderungan opini masyarakat lebih banyak positif mengenai peluncuran macbook model terbaru.

5.2. Saran

Dari penelitian yang telah dilakukan, dibuat saran untuk kedepannya agar dapat ditambahkan operator lain untuk mendapatkan hasil yang lebih akurat lagi, dan perlu dilakukan pengujian dengan algoritma lain untuk membandingkan tingkat akurasi. Dataset yang saat ini digunakan sebanyak 1015 data dan untuk kedepannya dapat ditambah lagi jumlah *record* datanya untuk mendapatkan hasil yang lebih baik lagi.

Daftar Pustaka

- [1] C. Ahmadi and I. G. R. P. Winata, "ANALISIS THROUGHPUT PENGIRIMAN DATA PADA JARINGAN WIRELESS DENGAN METODE QUEUE TREE," *JST (Jurnal Sains dan Teknologi)*, vol. 10, no. 1, pp. 112–116, May 2021, doi: 10.23887/jstundiksha.v10i1.30269.
- [2] T. Krisdiyanto, E. Maricha, and O. Nurharyanto, "Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naïve Bayes Clasifiers," *Jurnal CoreIT*, vol. 7, no. 1, 2021.
- [3] C. Mohi, H. Annur, and R. Pakaya, "Analisis Sentimen pada Tweets Divisi Humas Polri Dengan Metode Naive Bayes Classifier," *Jurnal Ilmiah Ilmu Komputer Banthayo Lo Komputer*, vol. 2, no. 1, pp. 34–43, May 2023, doi: 10.37195/balok.v2i1.509.
- [4] M. Winda, P. Sistem, I. Sekolah, T. Manajemen Informatika, D. Komputer, and N. Mandiri, "INTI NUSA MANDIRI ANALISIS SENTIMEN OPINI PUBLIK MENGENAI SARANA DAN TRANSPORTASI MUDIK TAHUN 2019 PADA TWITTER MENGGUNAKAN ALGORITMA NAÏVE BAYES, NEURAL NETWORK, KNN DAN SVM", [Online]. Available: <http://www.nusamandiri.ac.id>
- [5] B. Mas Pintoko and K. Muslim, "Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode Naïve Bayes Classifier."
- [6] D. Anjas Ramadhan and E. Budi Setiawan SSI, "ANALISIS SENTIMEN PROGRAM ACARA DI SCTV PADA TWITTER MENGGUNAKAN METODE NAIVE BAYES DAN SUPPORT VECTOR MACHINE."
- [7] N. Hardi *et al.*, "SISTEMASI: Jurnal Sistem Informasi Analisis Sentimen Physical Distancing pada Twitter Menggunakan Text Mining dengan Algoritma Naive Bayes Classifier." [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [8] I. Fahrur Rozi, A. Taufika Firdausi, and K. Islamiyah, "ANALISIS SENTIMEN PADA TWITTER MENGENAI PASCA BENCANA MENGGUNAKAN METODE NAÏVE BAYES DENGAN FITUR N-GRAM," *Jurnal Informatika Polinema*, vol. 6, no. 2, pp. 33–39, Mar. 2020, doi: 10.33795/jip.v6i2.316.
- [9] L. A. Andika, P. A. N. Azizah, and R. Respatiwan, "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*, vol. 2, no. 1, p. 34, Jul. 2019, doi: 10.13057/ijas.v2i1.29998.
- [10] W. Yulita, "Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier," *Jurnal Data Mining dan Sistem Informasi*, vol. 2, no. 2, p. 1, Aug. 2021, doi: 10.33365/jdmsi.v2i2.1344.
- [11] F. V. Sari and A. Wibowo, "ANALISIS SENTIMEN PELANGGAN TOKO ONLINE JD.ID MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER BERBASIS KONVERSI IKON EMOSI," *Jurnal SIMETRIS*, vol. 10, no. 2, 2019.