

KLASIFIKASI SENTIMEN PADA APLIKASI *VIEW RANGER* BERDASARKAN PENDAPAT PENGGUNA MENGGUNAKAN METODE *NAIVE BAYES*

Adrian Wijaya¹, Catur Nugrahaeni Puspita Dewi²
Informatika

Universitas Pembangunan Nasional Veteran Jakarta
Jl. RS. Fatmawati, Pondok Labu, Jakarta Selatan, DKI Jakarta, 12450, Indonesia
adrianwijaya@upnvj.ac.id¹, catur.nugrahaeni@upnvj.ac.id²

Abstrak. *View Ranger* merupakan aplikasi yang sangat berpengaruh bagi seseorang yang memiliki hobi mendaki gunung, bersepeda maupun *travelling*. Aplikasi ini mempunyai fitur yang canggih yaitu GPS *offline*. Pada tanggal 26 November 2021 berdasarkan data pada *Google Play* aplikasi *View Ranger* memiliki 45 ribu lebih ulasan yang diberikan oleh pengguna, memiliki *rating* 4.6 serta telah diunduh sebanyak 5 juta unduhan. Penelitian kali ini akan melakukan klasifikasi sentimen pada aplikasi *view ranger* berdasarkan pendapat pengguna atau ulasan menggunakan metode *naive bayes*. Pengambilan data dilakukan melalui *google play scraper* dengan menggunakan *google colab*. Hasil evaluasi klasifikasi menggunakan metode *naive bayes* yang menggunakan 1337 kata memperoleh nilai akurasi yaitu 0.8701 (87%), nilai *recall* yaitu 0.9318 (93%), nilai *precision* yaitu 0.8541 (85%), dan nilai *specificity* yaitu 0.7878 (78%). Dari hasil yang didapatkan dapat dikatakan bahwa performa metode *naive bayes* terbilang tinggi dalam tingkat akurasi, *recall*, *precision* serta *specificity* nya.

Kata Kunci: *Naive Bayes*, *View Ranger*, Klasifikasi Sentimen.

1. Pendahuluan

Seiring perkembangan teknologi di Indonesia, saat ini banyak memunculkan berbagai aplikasi – aplikasi inovatif untuk memudahkan segala macam keperluan manusia. Hal tersebut termasuk ke dalam berbagai macam jenis keperluan baik primer maupun sekunder, contoh dari keperluan sekunder seperti hal nya hiburan, rekreasi, hobi, dll. Salah satu perkembangan teknologi di bidang hobi yang inovatif yaitu aplikasi *View Ranger*.

View Ranger merupakan aplikasi yang sangat berpengaruh bagi seseorang yang memiliki hobi mendaki gunung, bersepeda maupun *travelling*. Aplikasi ini mempunyai fitur yang canggih yaitu GPS *offline*. Fitur ini sangat memudahkan seseorang dalam menggunakan aplikasi ini dalam keadaan ponsel online maupun offline [11]. Hadirnya aplikasi *View Ranger* juga memunculkan pro dan kontra antara pengguna GPS handheld dengan pengguna aplikasi ini, salah satu contoh kontra yang ditimbulkan yaitu mengenai tingkat keakuratan menangkap sinyal GPS yang dimiliki oleh aplikasi ini. Peneliti memilih aplikasi ini untuk penelitian karena aplikasi ini sudah populer di luar negeri namun belum begitu populer di Indonesia, hanya dibeberapa komunitas saja.

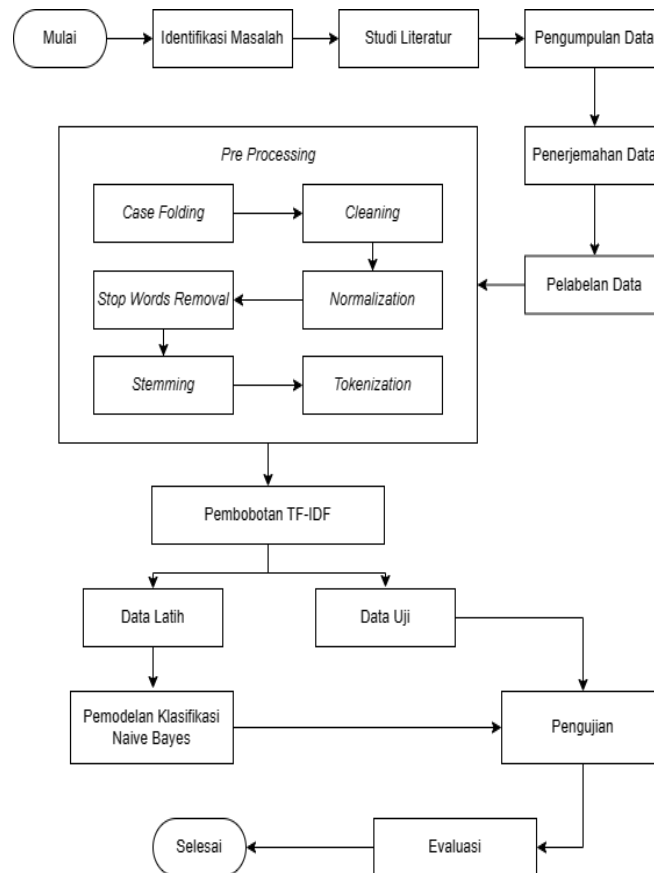
Pada tanggal 26 November 2021 berdasarkan data pada *Google Play* aplikasi *View Ranger* memiliki 45 ribu lebih ulasan yang diberikan oleh pengguna, memiliki *rating* 4.6 serta telah diunduh sebanyak 5 juta unduhan. Pengguna aplikasi *View Ranger* dapat memberikan tanggapan atau ulasan mereka terhadap aplikasi ini melalui komentar di *Google Play*. Dari ulasan – ulasan tersebut terdapat ulasan negative dan positive, dengan adanya penilaian dari pengguna aplikasi ini dapat menjadi bahan evaluasi untuk aplikasi *View Ranger* menggunakan analisis sentimen.

Penelitian sebelumnya yang dilakukan oleh Apriani et.al. mengenai analisis sentiment dengan *Naive Bayes* terhadap komentar aplikasi Tokopedia menghasilkan hasil dari pengujian yaitu nilai akurasi sebesar 97,13% dengan nilai *precision* 1 dan nilai AUC sebesar 0,980. Persentase yang dihasilkan untuk sentimen negative sebesar 63,53% sedangkan sentimen positive mendapatkan nilai sebesar 36,47% [1]. Berdasarkan peneliti Salehudin et.al. yang melakukan penelitian tentang analisis sentimen aplikasi dompet digital di era 4.0 pada masa pandemi covid-19 di *Play Store* menggunakan algoritma *Naive Bayes Classifier* menghasilkan hasil dari pengujian pada aplikasi Dana yaitu memiliki nilai akurasi sebesar 99,35% dengan sentimen positive berjumlah 182 ulasan [3].

Berdasarkan peneliti Yulita et.al yang melakukan penelitian tentang analisis sentimen terhadap opini masyarakat tentang vaksin covid-19 menggunakan algoritma naive bayes classifier menghasilkan hasil dari pengujian dari pendapat masyarakat memiliki tingkat akurasi yang tinggi yaitu sebesar 93% [12].

Melihat peneliti sebelumnya, peneliti ingin membuat penelitian yaitu klasifikasi sentimen pada aplikasi View Ranger berdasarkan pendapat pengguna menggunakan metode Naive Bayes guna mengetahui tingkat akurasi pada aplikasi ini yang akan dibagi menjadi dua bagian yaitu ulasan positive dan negative.

2. Metode Penelitian



Gambar. 1. Flowchart tahapan penelitian

2.1 Identifikasi Masalah

Mengidentifikasi permasalahan yang terdapat pada pengguna aplikasi *View Ranger* melalui komentar atau pendapat yang diberikan di *Google Play*.

2.2 Studi Literatur

Studi Literatur yaitu suatu proses yang dilakukan oleh peneliti ketika ingin memulai penelitian yaitu mencari sumber informasi serta data – data yang diperlukan berkaitan dengan penelitian yang akan dilakukan. Proses pencarian data bisa didapat dari membaca buku, jurnal serta internet.

2.3 Pengumpulan Data

Pengumpulan data pada penelitian ini menggunakan *google colab* dan *google play scraper*. Proses ini merupakan langkah pertama dalam penelitian. Untuk melakukan *scraping* data menggunakan *google play scraper* melalui *google colab*. Keluaran dari proses ini yaitu berupa data mentah yang akan digunakan dalam bahasa Inggris.

2.4 Penerjemahan Data

Penerjemahan data pada penelitian ini menggunakan *google document* dari *google translate*. Proses ini merupakan langkah kedua setelah pengumpulan data. Keluaran dari proses ini yaitu berupa data mentah yang akan digunakan dalam bahasa Indonesia.

2.5 Pelabelan Data

Setelah data dikumpulkan, data tersebut akan masuk kedalam proses pelabelan data. Pelabelan data ini bertujuan untuk memisahkan antara ulasan yang berupa sentimen *negative* dengan ulasan yang berupa sentimen *positive*. Pelabelan data pada proses ini dilakukan secara manual yaitu melalui 3 annotator. Indikator dalam pelabelan yaitu untuk ulasan yang menggambarkan kesenangan maupun hal – hal yang bersifat positif akan dimasukkan kedalam kelas label sentimen positif dan untuk ulasan yang menggambarkan sebaliknya yaitu berupa kekecewaan atau hal – hal yang bersifat negatif akan dimasukkan kedalam kelas label sentimen negatif. Untuk hasil akhir pelabelan diambil sesuai label terbanyak dari ketiga annotator. Pelabelan data yang dilakukan secara manual ditakutkan dapat mengakibatkan kesalahan, oleh karena itu peneliti memakai metode fleiss kappa untuk mengetahui data yang digunakan konsisten atau tidak.

2.6 Pre Processing

Proses *pre processing* sangat diperlukan karena dari seluruh data yang diambil untuk diuji, diantaranya sering kali terdapat berbagai permasalahan yang membuat data tersebut tidak sesuai dengan format yang dibutuhkan oleh sistem. Oleh karena itu pada tahapan *pre processing* ini digunakan untuk menghilangkan berbagai permasalahan pada data agar data yang dibutuhkan sesuai oleh format yang dibutuhkan oleh *system*.

2.6.1. Case Folding

Tahapan ini disebut *case folding* karena terjadinya perubahan semua huruf besar atau kapital yang terdapat pada data menjadi huruf kecil.

2.6.2. Cleaning

Cleaning adalah proses pembersihan tanda baca, seperti tanda seru (!), tanda tanya (?), titik (.) atau tanda baca yang lainnya.

2.6.3. Normalization

Normalization adalah proses normalisasi suatu bahasa. Selama proses ini, jika ada kata yang tidak baku akan dinormalkan kedalam bentuk yang baku sesuai kaidah yang tertuai pada Kamus Besar Bahasa Indonesia (KBBI).

2.6.4. Stop Words Removal

Pada tahapan ini dilakukan penghilangan kata yang sama sekali tidak memiliki arti yang berpengaruh kedalam kata sentimen seperti kata penghubung, waktu.

2.6.5. Stemming

Pada tahapan ini yaitu proses *stemming* dilakukan perubahan setiap kata – kata imbuhan menjadi kata dasar.

2.6.6. Tokenization

Pada tahapan tokenisasi ini dilakukan pemisahan setiap kata, kata tersebut dipisahkan ketika adanya spasi antar kata.

2.7 Pembobotan TF – IDF

Pada tahap ini dilakukan pembobotan TF-IDF terhadap data yang telah diolah terlebih dahulu. Tahapan ini bertujuan untuk memberikan nilai pada setiap kata untuk mempermudah proses klasifikasi *Naive Bayes Classifier*.

2.8 Pembagian Data

Setelah melalui proses *pre processing*, data keseluruhan yang dimiliki dibagi menjadi dua bagian yaitu berupa data latih dan data uji. Pembagian data yaitu sebesar 80% data latih dan data uji sebesar 20%. Model dibentuk menggunakan data latih yang telah melalui proses pelabelan dan *pre processing*. Data uji digunakan untuk klasifikasi menggunakan model yang didapat dari data latih yang telah dibuat.

2.8 Analisis dan Hasil

Proses pengklasifikasian data pada tahapan analisis yaitu menggunakan algoritma *Naive Bayes*. Data yang dihasilkan berupa jumlah dari banyaknya sentimen *positive* maupun *negative*. Hasil yang didapatkan dapat dipresentasikan dalam bentuk tabel agar lebih mudah dipahami. Setelah mendapatkan hasil dari proses klasifikasi maka akan dilakukan evaluasi menggunakan metode *confusion matrix*. Tahap ini dilakukan untuk mengukur performa model yang sudah dibangun dimana hasil dari evaluasi berupa nilai *accuracy*, *recall*, *precisison*, dan *specificity*.

3. Hasil dan Pembahasan

Pada penelitian kali ini peneliti menggunakan data yang diambil dari ulasan dari *google play* pada aplikasi *View Ranger*. Untuk pengambilan data, peneliti menggunakan metode *scraping* data dengan aplikasi *google colab*. Data yang di *scraping* yaitu berupa nama serta ulasan komentar yang menggunakan aplikasi *View Ranger*.

Berikut adalah hasil data mentah dari *scraping* melalui *google play scraper* yang telah diterjemahkan kedalam bahasa Indonesia.

1	isi
2	Pencurian! Saya membayar untuk aplikasi VR dan peta OS yang saya kehilangan aksesnya - meskipun saya membayar untuk aplikasi t
3	Harus jelas dari banyak komentar bahwa pengguna viewranger menyukai aplikasi apa adanya. Saya tidak memerlukan perangkat tam
4	Saya membeli 8 peta OS atau lebih dan membayar masing-masing £7,99. Saya kehilangan semuanya saat aplikasi ini diakuisisi oleh O
5	apakah itu navigasi kapal?, versi lama apakah tersedia?
6	Sangat mengecewakan melihat VR menyatakan perubahan itu.
7	Versi lama lebih baik dari yang baru Hapus banyak fitur di versi baru Tidak layak untuk dibayar karena itu bukan aplikasi yang pantas r
8	Mengapa melakukan ini. Aplikasi baru ini tidak sebagus itu
9	Sangat bagus untuk menganalisis jalan-jalan Anda
988	Aplikasi brilian, bagus untuk melacak lari, berjalan, dll, dan cadangan tambahan untuk navigasi
989	Aplikasi ini telah membantu saya menemukan begitu banyak jalan tersembunyi di area lokal saya, dan untuk Â£25 setahun untuk me
990	Hal terbaik untuk hiking, menandai jalur untuk berbagi. Sangat menyarankan menggabungkan fungsi multi-bahasa. Cina Spanyol Jeps
991	Salah satu aplikasi hiking terbaik. Bagus untuk mereka yang mau membayar 4 langganan OS yang terjangkau. Secara pribadi saya lebih
992	Alat luar biasa untuk memetakan jalan dan merekam trek gpx untuk OpenStreetMap
993	Normal
994	Itu hanya bekerja. Sebagian besar. Terkadang sulit untuk dinavigasi. Umumnya bekerja dengan baik.
995	Membuat pendakian Anda lebih menyenangkan terutama jika cuaca tidak terlalu bagus seperti yang saya alami, membuat saya kemb
996	Besar. Tidak pernah gagal. Alat pengukur jarak yang sangat akurat.
997	Terkadang bisa sedikit lambat. Tapi id pengalaman Anda dalam membaca peta kadang-kadang membutuhkan waktu untuk menemukan
998	Ini benar-benar relabe dan luar biasa
999	Aplikasi favorit untuk melacak kenaikan dan lari.
1000	Aplikasi hebat dengan banyak jalan yang berbeda
1001	

Gambar. 2. Data Mentah

Setelah itu masuk ke tahap pelabelan data. Pada tahap pelabelan data, peneliti dibantu oleh 3 annotator yang akan melakukan pelabelan secara manual. Annotator akan melabeli data yang memiliki sentimen *negative* dengan kata ganti “0” sedangkan untuk data yang memiliki sentimen *positive* dengan kata ganti “1”. Berikut hasil pelabelan yang telah dilakukan oleh ketiga annotator.

Tabel 1. Hasil Pelabelan Data (*Sample Data Latih*)

Ulasan	Annotator 1	Annotator 2	Annotator 3	Label Hasil
Saya sangat terkesan dgn aplikasi ini dan akan merekomendasikannya. Saya tdk akan pergi berjalan tanpa itu.	1	1	1	1
Saya suka menggunakan viewranger utk jalan-jalan saya - teman-teman saya juga selalu terkesan	1	1	1	1
Lebih lambat dari siput kering! Tidak terinstal.	0	0	0	0

Sebelum masuk kedalam tahapan model klasifikasi data masuk kedalam tahapan *pre processing* guna membersihkan data – data yang masih mentah menjadi data bersih agar lebih mudah saat masuk kedalam tahap klasifikasi.

Tabel 2. Hasil *Pre Processing* Data

Data Sesudah <i>Pre Processing</i>	
D1	['kesan', 'aplikasi', 'rekomendasi', 'pergi', 'jalan']
D2	['suka', 'guna', 'viewranger', 'jalan', 'jalan', 'teman', 'teman', 'kesan']
D3	['lambat', 'siput', 'kering', 'instal']

Tahapan ini dilakukan setelah data sudah selesai melalui proses *pre processing*. TF-IDF atau *Term Frequency-Inverse Document Frequency* adalah metode yang digunakan pada tahap pembobotan kata pada penelitian ini. Sesuai namanya, metode ini berfungsi untuk menghitung frekuensi dari data atau banyaknya kata yang muncul pada suatu data.

Tabel 3. Hasil Pembobotan TF - IDF

Term	TF			DF	IDF	TF-IDF		
	D1	D2	D3			D1	D2	D3
kesan	1	1	0	2	0,1760913	0,1760913	0,1760913	0
aplikasi	1	0	0	1	0,4771213	0,4771213	0	0
rekomendasi	1	0	0	1	0,4771213	0,4771213	0	0
pergi	1	0	0	1	0,4771213	0,4771213	0	0
jalan	1	2	0	3	0	0	0	0
suka	0	1	0	1	0,4771213	0	0,4771213	0
guna	0	1	0	1	0,4771213	0	0,4771213	0
viewranger	0	1	0	1	0,4771213	0	0,4771213	0
teman	0	2	0	2	0,1760913	0	0,3521825	0
lambat	0	0	1	1	0,4771213	0	0	0,4771213
siput	0	0	1	1	0,4771213	0	0	0,4771213
kering	0	0	1	1	0,4771213	0	0	0,4771213
instal	0	0	1	1	0,4771213	0	0	0,4771213

Pada tahapan klasifikasi, data – data yang telah melalui tahapan *pre processing* dan tahapan pembobotan TF – IDF akan diklasifikasikan menggunakan model *Naive Bayes*. Sebelum memasuki tahapan klasifikasi data akan memasuki tahapan pembagian data terlebih dahulu.

Data yang berjumlah 770 akan dipisah secara acak menjadi dua bagian yaitu menjadi data latih dan data uji. Pembagian data yaitu sebesar 80% untuk data latih dan 20% uuntuk data uji. Pembagian data dilakukan menggunakan library train test split pada python.

Tabel 4. Hasil Pembagian Data

Pembagian Data	Positif	Negatif	Jumlah	Jumlah Kata
Data Latih	372	244	616	
Data Uji	88	66	154	1337
Total	460	310	770	

Setelah itu dilakukan pemodelan naive bayes pada data latih.

Tabel 5. Nilai Probabilitas Kata *Sample Data*

Term	TF-IDF			P(X Positive)	P(X Negative)
	D1	D2	D3		
kesan	0,176091	0,176091	0	0,061462842	0,058823529

aplikasi	0,477121	0	0	0,067141875	0,058823529
rekomendasi	0,477121	0	0	0,067141875	0,058823529
pergi	0,477121	0	0	0,067141875	0,058823529
jalan	0	0	0	0,045454545	0,058823529
suka	0	0,477121	0	0,067141875	0,058823529
guna	0	0,477121	0	0,067141875	0,058823529
viewranger	0	0,477121	0	0,067141875	0,058823529
teman	0	0,352183	0	0,061462842	0,058823529
lambat	0	0	0,477121	0,045454545	0,086889486
siput	0	0	0,477121	0,045454545	0,086889486
kering	0	0	0,477121	0,045454545	0,086889486
instal	0	0	0,477121	0,045454545	0,086889486

Setelah dilakukan nya proses perhitungan pada data latih. Selanjutnya akan dilakukan proses perhitungan pada data uji berdasarkan hasil yang didapat oleh perhitungan data latih.

Berikut contoh sample data uji

Tabel 6. *Sample Data Uji*

Data Mentah	Data Setelah <i>Pre Processing</i>
Suka. ViewRanger adalah aplikasi yang sangat bagus	['suka', 'viewranger', 'aplikasi', 'bagus']

Setelah dilakukannya *pre processing* pada data uji, lalu data uji akan melalui tahapan pembobotan TF-IDF. Berikut tabel perhitungannya.

Tabel 7. *Pembobotan TF – IDF Sample Data Uji*

Term	TF	DF	IDF	TF-IDF
suka	1	1	0,47712	0,47712

viewranger	1	1	0,47712	0,47712
aplikasi	1	1	0,47712	0,47712
bagus	1	1	0	0

Pada tabel merupakan hasil pembobotan kata TF-IDF dari data *sample* uji. Kata “bagus” bernilai “0” karena kata tersebut tidak ditemukan pada data latih. Selanjutnya dilakukan perhitungan probabilitas dari setiap kata dari *sample* data uji.

$$P(W_i|p) = \frac{\text{count}(w_{ij,p})+1}{|p|+|V|*1} = \frac{\text{count}(w_{ij,p})+1}{|p|+|V|} \quad (1)$$

Keterangan :

$P(W_i|p)$ = Peluang munculnya kata W_i pada kelas p.

$\text{count}(w_{ij,p})$ = Jumlah munculnya kata W_i pada kelas p.

$|p|$ = Total keseluruhan kata pada kelas p.

$|V|$ = Total keseluruhan kata (*term*).

1. Probabilitas kata pada kelas label positive

$$P(\text{suka}|\text{positif}) = \frac{\text{count}(w_{\text{suka},\text{positif}}) + 1}{9 + 13} = \frac{0.47712 + 1}{9 + 13} = 0.0671418$$

$$P(\text{viewranger}|\text{positif}) = \frac{\text{count}(w_{\text{viewranger},\text{positif}}) + 1}{13 + 13} = \frac{0.47712 + 1}{9 + 13} = 0.0671418$$

$$P(\text{aplikasi}|\text{positif}) = \frac{\text{count}(w_{\text{aplikasi},\text{positif}}) + 1}{9 + 13} = \frac{0.47712 + 1}{9 + 13} = 0.0671418$$

$$P(\text{bagus}|\text{positif}) = \frac{\text{count}(w_{\text{bagus},\text{positif}}) + 1}{9 + 13} = \frac{0 + 1}{9 + 13} = 0.0454545$$

2. Probabilitas kata pada kelas label negative

$$P(\text{suka}|\text{negatif}) = \frac{\text{count}(w_{\text{suka},\text{negatif}}) + 1}{4 + 13} = \frac{0 + 1}{4 + 13} = 0.0588235$$

$$P(\text{viewranger}|\text{negatif}) = \frac{\text{count}(w_{\text{viewranger},\text{negatif}}) + 1}{4 + 13} = \frac{0 + 1}{4 + 13} = 0.0588235$$

$$P(\text{aplikasi}|\text{negatif}) = \frac{\text{count}(w_{\text{aplikasi},\text{negatif}}) + 1}{4 + 13} = \frac{0 + 1}{4 + 13} = 0.0588235$$

$$P(\text{bagus}|\text{negatif}) = \frac{\text{count}(w_{\text{bagus},\text{negatif}}) + 1}{4 + 13} = \frac{0 + 1}{4 + 13} = 0.0588235$$

Setelah menghitung probabilitas setiap kata pada data uji, untuk menentukan data uji tersebut masuk ke dalam kelas label *positive* atau *negative* dilakukannya perhitungan probabilitas masing – masing kelas.

Perhitungan probabilitas kelas label *positive* dan *negative* pada data uji

$$P(\text{positif}|\text{data uji}) = 0.66 * 0.0671418 * 0.0671418 * 0.0671418 * 0.0454545 = 9,08029E-06$$

$$P(\text{negatif}|\text{data uji}) = 0.33 * 0.0588235 * 0.0588235 * 0.0588235 * 0.0588235 = 3,95109E-06$$

Penentuan kelas label pada data uji diambil dari nilai terbesar probabilitas yang dihasilkan. Dari data perhitungan

di atas, nilai $P(\text{positif}|\text{data uji}) > P(\text{negatif}|\text{data uji})$. Oleh karena itu, data uji masuk kedalam kelas label *positive*.

Setelah melakukan perhitungan pada data latih dan perhitungan serta penentuan kelas label pada data uji. Tahapan akhir yaitu evaluasi yang bertujuan untuk menampilkan dari model klasifikasi *naive bayes*.

Data yang didapatkan dari perhitungan dari seluruh data uji yang berjumlah 154 yaitu data positif yang benar (TP) berjumlah 82 data, data negatif yang benar (TN) berjumlah 52 data, data positif yang salah (FP) berjumlah 14 data, data negatif yang salah (FN) berjumlah 6 data.

Tabel 8. *Confusion Matrix*

		Benar	
		Positif	Negatif
Prediksi	Positif	82 (TP)	14 (FP)
	Negatif	6 (FN)	52 (TN)

Perhitungan model naive bayes dihitung menggunakan persamaan (2), (3), (4) dan (5). Nilai yang dihitung yaitu nilai akurasi, recall, precision, dan specificity.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (5)$$

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FN + FP} = \frac{82 + 52}{82 + 52 + 6 + 14} = \frac{134}{154} = 0.8701$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{82}{82 + 6} = \frac{82}{88} = 0.9318$$

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{82}{82 + 14} = \frac{82}{96} = 0.8541$$

$$\text{Specificity} = \frac{TN}{TN + FP} = \frac{52}{52 + 14} = \frac{52}{66} = 0.7878$$

Hasil perhitungan klasifikasi model *naive bayes* berdasarkan perhitungan diatas mendapatkan nilai akurasi sebesar 0.8701 (87%), nilai *recall* sebesar 0.9318 (93%), nilai *precision* sebesar 0.8541 (85%), dan nilai *specificity* sebesar 0.7878 (78%).

4. Kesimpulan dan Saran

4.1 Kesimpulan

Berdasarkan penelitian yang dilakukan yaitu “klasifikasi sentimen pada aplikasi *view ranger* berdasarkan pendapat pengguna menggunakan metode *naive bayes*”, mendapatkan beberapa kesimpulan sebagai berikut.

1. Penelitian ini menggunakan metode *naive bayes*. Data diambil dari ulasan *google play* berjumlah 1000 data, kemudian data diseleksi kembali secara manual menjadi 770 data yang digunakan. Terdapat 460 data yang mempunyai nilai sentimen *positive* dan 310 data yang mempunyai nilai sentimen *negative*. Data tersebut melalui proses pelabelan secara manual oleh 3 annotator, setelah itu melalui tahapan *pre processing* untuk membersihkan data – data sebelum masuk kedalam tahap klasifikasi yang menggunakan metode *naive bayes*. Data yang telah bersih memasuki tahapan pembobotan TF – IDF . Setelah melalui 3 tahapan tersebut data dipisah menjadi 2 bagian yaitu berupa data latih dan data uji, 80% untuk data latih dan 20% untuk data uji. Data latih berjumlah 616 data sedangkan data uji berjumlah 154 data dengan total kata yang didapatkan dari 770 data yaitu 1337 kata. Lalu setelah data dibagi menjadi dua, data tersebut diklasifikasikan menggunakan metode *naive bayes*. Data juga divisualisasikan untuk melihat kata – kata dominan yang muncul baik untuk sentimen positif maupun negatif dalam bentuk *wordcloud*.
2. Hasil evaluasi klasifikasi menggunakan metode *naive bayes* yang menggunakan 1337 kata memperoleh nilai akurasi sebesar 0.8701 (87%), nilai *recall* sebesar 0.9318 (93%), nilai *precision* sebesar 0.8541 (85%), dan nilai *specificity* sebesar 0.7878 (78%). Dari hasil yang didapatkan dapat dikatakan bahwa performa metode *naive bayes* terbilang tinggi dalam tingkat akurasi, *recall*, *precision* serta *specificity* nya.

4.2 Saran

Terdapat beberapa saran yang dapat diberikan oleh peneliti setelah melakukan penelitian kali ini.

1. Data yang akan digunakan lebih banyak untuk penelitian.
2. Dapat menggunakan metode lainnya dalam melakukan proses klasifikasi seperti *Support Vector Machine* (SVM), *Chi Square*, *Particle Swarm Optimization* (PSO), dan lain-lain.
3. Dalam menerjemahkan data dapat menggunakan translator yang lebih mumpuni dan akurat.

5. Referensi

- [1] Apriani, R. et.al., 2019. ANALISIS SENTIMEN DENGAN NAÏVE BAYES TERHADAP KOMENTAR APLIKASI TOKOPEDIA. *Jurnal Rekayasa Teknologi Nusa Putra*, Volume 6, Nomor 1, Halaman 54 – 62.
- [2] Arie, F. et.al., 2019. PENGARUH KATA CASHBACK TERHADAP PENINGKATAN PENJUALAN MENGGUNAKAN DATA MINING. *Jurnal Ilmiah Manajemen Informatika dan Komputer*, Volume 3, Nomor 2, Halaman 1 – 5.
- [3] Basryah, Egi Salehudin.2021. ANALISIS SENTIMEN APLIKASI DOMPET DIGITAL DI ERA 4.0 PADA MASA PANDEMI COVID-19 DI PLAY STOREMENGUNAKAN ALGORITMA NAIVE BAYESCLASSIFIER. *Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra*.
- [4] Google Play,2021.“ViewRangerJejak&Peta”,<https://play.google.com/store/apps/details?id=com.augmentra.viewranger.android&hl=in&gl=US>, diakses pada 22 November 2021.
- [5] Indrayuni, E., 2018. Komparasi Algoritma Naive Bayes dan Support Vector Machine untuk Analisa Sentimen Riview Film. *Jurnal Pilar Nusa Mandiri*, XIV(2), Halaman 175.
- [6] Luqyana, W. A., Cholissodin, I., & Perdana, R. S. (2018). Analisis Sentimen Cyberbullying Pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 2(11), 4704–4713.
- [7] Pramadani, R. et.al., 2020. PENERAPAN TEXT MINING DALAM MENGANALISIS KEPERIBADIAN PENGGUNA MEDIA SOSIAL. *Jurnal Teknik Informatika Musirawas*, Volume 5, Nomor 01.

- [8] Sabrani, A., Wedashwara W., I. G. W., & Bimantoro, F. (2020). Multinomial Naïve Bayes untuk Klasifikasi Artikel Online tentang Gempa di Indonesia. *Jurnal Teknologi Informasi, Komputer, Dan Aplikasinya (JTika)* , 2(1), 89–100. <https://doi.org/10.29303/jtika.v2i1.87>
- [9] Saritas, M. M. & Yasar, A., 2019. Performance Analysis of ANN and Naive Bayes Classification Algorithm. *International Journal of Intelligent Systems and Applications in Engineering*, VII(2), Halaman 88-91.
- [10] Srinivasa, K, G Siddesh, and H Srinidi. 2018. Network Data Analytics: A Hands-On Approach for Application Development.
- [11] Yoiyok. 2019. “Aplikasi Android GPS Offline (View Ranger) Untuk Mendaki Gunung”, <https://www.santridanalam.com/2019/01/aplikasi-android-gps-offline-viewranger.html>, diakses pada 22 November 2021.
- [12] Yutika, C. H., Adiwijaya, A., & Faraby, S. Al. (2021). Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes. *Jurnal Media Informatika Budidarma*, 5(2), 422–430. <https://doi.org/10.30865/mib.v5i2.2845>