

---

## Analisa Sentimen Terhadap Review *Fintech* Dengan Metode Naive Bayes Classifier Dan K- Nearest Neighbor

Surohman<sup>1</sup>, Sopian Aji<sup>2</sup>, Rousyati<sup>3</sup>, Fanny Fatma Wati<sup>4</sup>

<sup>1</sup> Magister Ilmu Komputer, STMIK Nusa Mandiri Jakarta

<sup>2</sup> Sistem Informasi, STMIK Nusa Mandiri Jakarta

<sup>3</sup> Sistem Informasi, Universitas Bina Sarana Informatika

<sup>3</sup> Sistem Informasi Akuntansi, Universitas Bina Sarana Informatika

Indonesia

E-mail:surohman072016@gmail.com, sopian.sop@nusamandiri.ac.id, rousyati.rou@bsi.ac.id,  
fanny.ffw@bsi.ac.id

**Abstract** - Based on observations made by researchers, *FinTech* which is common and is being used by the people of Indonesia is the first generation for payment, purchase of goods and services, one of which is the Fund Application. In choosing the application to be used, it is usually considered, comfort, safety, accuracy of the transaction, comfort, and various promotions. But some users are hesitant in using the application because of some reviews from the application that show positive and negative ratings. With the number of reviews paid in the comments column provided by the Google Play Store in the Funds Application needed to classify the reviews given as positive or negative. The research used was an experimental method using the Naive Bayes Classifier Algorithm and K-Nearest Neighbor. Regarding testing on the Fund Application has the best testing value of 84.76%. Then it can be considered a review that can be approved from positive reviews by other users and if there are negative reviews it will be input to the company to develop and develop the product.

**Keywords:** *Sentiment Analysis, Fintech, Fund Application*

**Abstract** - Berdasarkan pengamatan yang dilakukan oleh peneliti, *FinTech* yang umum dan sedang marak digunakan oleh masyarakat Indonesia terutama generasi milineal untuk pembayaran, pembelian barang dan jasa salah satunya adalah Aplikasi Dana. Dalam memilih aplikasi yang akan digunakan biasanya mempertimbangkan, kenyamanan, keamanan, ketepatan transaksi, kemudahan, dan banyaknya promosi. Tetapi beberapa pengguna terkadang masih ragu dalam menggunakan suatu aplikasi karena beberapa review dari aplikasi yang menunjukkan ulasan positif maupun negatif. Dengan banyaknya review yang ditampilkan pada kolom komentar yang telah disediakan oleh Google Play Store di Aplikasi Dana dibutuhkan analisa untuk mengklasifikasi ulasan yang diberikan termasuk positif atau negatif. Penelitian yang dipakai adalah metode eksperimen dengan menggunakan Algoritma Naive Bayes Classifier dan K-Nearest Neighbor. Hasilnya pengujian pada Aplikasi Dana memiliki nilai akurasi terbaik sebesar 84,76%. Maka dapat menyimpulkan review tersebut dapat meyakinkan dari review positif oleh pengguna lain dan jika terdapat review negatif maka akan menjadi masukan kepada

pihak perusahaan untuk lebih mengembangkan maupun memperbaiki produk tersebut.

**Keywords:** Analisa Sentimen, Fintech, Aplikasi Dana

## 1. Introduction

Penggunaan smartphone sebagai sarana telekomunikasi kian berkembang. Di Indonesia pengguna smartphone semakin banyak oleh sebab itu Indonesia mempunyai julukan sebagai raksasa teknologi digital di kancah Asia. Jumlah pengguna aktif smartphone di Indonesia hampir mencapai setengah dari jumlah keseluruhan penduduk, jumlah tersebut berdasarkan survei lembaga riset digital marketing. Dari banyaknya jumlah pengguna tersebut maka Indonesia menjadi negara pengguna smartphone terbesar ke empat di dunia setelah Cina dan yang menduduki tingkat pertama adalah Amerika. Penggunaan teknologi yang semakin meluas akan berdampak pada berkembangnya dunia bisnis, sesuai dengan riset yang disimpulkan oleh DBS Group (Kominfo, 2019). Di generasi milenials kemudahan teknologi membuat fenomena phubbing yaitu ketergantungan seseorang dengan smartphone. Berbagai manfaat yang dapat dilakukan dengan smartphone seperti aktivitas komunikasi, sosial media, pembayaran, transfer uang, pemesanan tiket, belanja online, dan masih banyak lagi

manfaat dari perkembangan teknologi (Hanika, 2015).

Selain penggunaan teknologi yang telah disebutkan, teknologi juga dapat digunakan sebagai financial technology (Fintech) yang saat ini semakin dikenal dan semakin mudah digunakan oleh masyarakat dalam melakukan sebagai salah satu metode pembayaran yang lebih praktis dan efektif. Fintech atau teknologi keuangan didefinisikan sebagai inovasi teknologi dalam layanan keuangan yang menghasilkan model-model bisnis, aplikasi, proses, produk dengan efek material yang berkaitan dengan layanan keuangan. Fintech saat ini membantu generasi milenial dalam membantu proses pelayanan produk, literasi keuangan, pembayaran, transfer secara digital. Otoritas Jasa Keuangan (OJK) telah mencatat 64 perusahaan fintech yang sudah terdaftar (Sugiarti, Diana, & Mawardi, 2019). Bank Indonesia memberi empat paparan mengenai fintech yaitu fintech sebagai Crowdfunding dan Peer to Peer (P2P) Lending, Market Agregator, Risk and Investment Management serta Paymet , Settlement, and Clearing (Suyanto & Kurniawan, 2019).

Berdasarkan pengamatan yang dilakukan oleh peneliti, FinTech yang umum dan sedang marak digunakan oleh masyarakat Indonesia terutama generasi milineal untuk pembayaran, pembelian barang dan jasa adalah Dana, Ovo dan Gopay. Dilihat dari Google Play Store Gopay telah didownload 50 juta lebih dengan 2 juta ulasan, Ovo telah didownload 10 juta lebih dengan 268 ribu lebih ulasan sedangkan Dana telah didownload 10 jt lebih dengan 321 ribu lebih ulasan. Dana dan Ovo saling bersaing. Ketiga aplikasi tersebut sedang gencar promosi, banyak kerjasama dengan pelaku bisnis serta pemakai dapat mengumpulkan banyak poin sehingga mendapat potongan pembayaran. Pada aplikasi Google Playstore sendiri menyediakan fitur komentar sebagai wadah untuk pengguna dalam mereview tentang aplikasi yang sudah di install dan digunakan pada smartphone. Manfaat fitur digunakan pengguna untuk mengungkapkan perasaan, emosi, kebanggaan, dan ulasan lain mengenai aplikasi. Ulasan yang diberikanpun terdiri dari berbabagai macam pendapat baik negatif maupun positif.

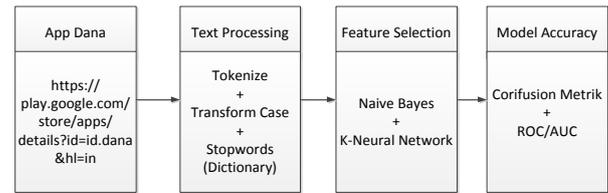
Dalam analisis sentimen klasifikasi dibagi menjadi tiga tingkatan, yaitu tingkat dokumen, tingkat kalimat, dan tingkat aspek. Dalam penelitian ini, analisis sentimen mengambil tingkat kalimat, dengan menggunakan metode Naïve Bayes Classifier dan K- Neural Network Analisa

Sentimen biasanya mengacu pada perasaan, emosi, sikap maupun pendapat. Karena yang digunakan data tekstual maka terdapat kebutuhan untuk menganalisa konsep mengekspresikan sentimen dan menghitung wawasan untuk menjelajahi bisnis. Analisa sentimen merupakan suatu proses untuk menentukan apakah ulasan yang diungkapkan cenderung ulasan positif, negatif atau netral . Opini mining tidak memperhatikan topik dari teks tersebut tetapi lebih fokus kepada ekspresi yang digambarkan dari teks opini tersebut. Hal ini menentukan review atau ulasan berkaitan dengan aplikasi Dana pada Google Play Store termasuk opini positif, negatif atau netral.

Penelitian ini akan menganalisa mengenai permasalahan yang berkaitan dengan beberapa ulasan tentang aplikasi Dana di Google Play Store serta menentukan hasil akurasi analisa sentimen yang dihasilkan algoritma Naïve Bayes dan K-Nearest Neighbors. Adapun manfaat dari penelitian ini adalah untuk membantu manajemen Aplikasi Dana mengenai opini positif atau negatif dari pengguna aplikasi serta dapat memberikan bukti secara empiris untuk teori yang berkaitan sehingga dapat dijadikan sumbangan pemikiran untuk pengembangan teori berikutnya.

## 2. Materials and Methods

Metode yang digunakan dalam penelitian ini menggunakan metode eksperimen. Penelitian yang dilakukan pada analisis sentimen review dengan menggunakan Algoritma Naive Bayes Classifier dan K-Nearest Neighbors pada aplikasi Dana menggunakan Aplikasi RapidMiner versi 9.2 untuk mengukur akurasi data eksperimen. Untuk melakukan eksperimen peneliti menggunakan data komentar review pada aplikasi Dana yang ada pada Google Playstore. Data yang di ambil sebanyak 232 data review yang terdiri dari 116 data review positif dan 116 data review negatif. Text processing yang peneliti gunakan Tokonize, Transform Case, Stopword (Dictionary). Pengujian yang dilakukan dengan memilih fitur seleksi Naive Bayes dan K-Nearest Neighbors yang bertujuan untuk mendapatkan nilai akurasi. Dari nilai akurasi persamaan model confusion matrix proporsi jumlah nilai prediksi didapatkan dan ROC Curve sehingga kita bisa melihat grafik ROC dengan nilai AUC (Area Under Curve). Metode penelitian eksperimen dilakukan dalam penelitian dengan menggunakan kerangka pemikiran yang diusulkan dalam penelitian sebagai berikut:



Gambar. 1. Kerangka Pemikiran

### 2.1. Pengumpulan Data

Pada Gambar.1 menjelaskan bahwa kerangka pemikiran dalam penelitian ini dimulai dengan mengambil review pada aplikasi Dana yang ada pada playstore <https://play.google.com/store/apps/details?id=id.dana&hl=in> data yang digunakan adalah komentar pengguna aplikasi tersebut dengan jumlah 232 review terdiri dari 116 review positif dan 116 review negatif. Dari hasil pengumpulan data tersebut kemudian dipilah dan melakukan pengklasifikasian review negatif dan positif proses ini termasuk teknik preprocessing data.

Tabel I  
Dataset Penelitian

Nilai	Review
Positif	Ini aplikasi luar biasa membantu, pengisian gampang, transfer antar bank gampang banget dan gak ada potongn biaya, mau transaksi apapun ga pernah ada kendala, cuma sarannya tolong supaya untuk pengisian saldo bisa dilakukan di BRI juga. Soalnya saya pake BRI
Positif	Aplikasi yang sangat mudah digunakan,, dan recommended deh buat kalian yang mau mendapatkan diskon ataupun cashback dari DANA,, dan bagi kalian yang belum download aplkasi ini buruan download dan

	dapatkan voucher sebesar 25.000 buat kalian yang menggunakan REFERRAL dari aku Caranya mudah kok,, tingg...
Positif	puas bgt. bsnyak discount, terutama utk beli pulsa ini sumpah recomended bgt. tf antar bank lebih mudah dgn biayaya admin 0 rupiah alias gratis tis tiiiiissss. ga perlu repot <sup>2</sup> bawa duit cash, bayar ini itu tinggal scan jd bisa lebih less cash :) cuma sayang td di indomart ga bisa top up, dulu bisa ....
Positif	Bagus modernisasinya, meminimalisir jumlah uang kertas bisa menaikkan nilai rupiah. tampilan rapih semua oke, banyak voucher, dan jangan berharap gratis dan dapet uang dari DANA, karena DANA aplikasi pemudah bukan peng-gratis. terima kasihh Add kode dari saya yaa, bisa langsung dapet voucher 25.000 l...
Positif	Sangat membantu dan bermanfaat. Sedikit saran Mohon ditambahkan fitur seperti cetak struk, pay later dll. Saat ini DANA is the best sayang fiturnya masih harus dikembangkan lg. Mungkin masukan dari teman" DANA bisa dipertimbangkan.
Negarif	Aplikasi penipu. Voucher yang diberikan tidak bisa dipakai dengan alasan transaksi saya tidak tidak sesuai ketentuan promo. Memiliki 2 id di 1 device. Tapi sewaktu saya minta bukti, tidak bisa menunjukannya.
Negarif	Saya ada uang masuk dr bukalapak, ktk saya mau cairkan trnyta dana meminta saya utk verifikasi Ektp, dsini lah masalahnya foto bolak balik ditolak krn kualitas foto kuranh baik, saya sdh tlp kustumer servis dr kmrn katanya suruh kirim email dng memegang ktp, sdh

	saya kirim tp tdk direspon juga, APLIKASI NYUSAHIN, HATI2 UNTUK KAWAN2 JANGAN SAMPAI MENGALAMI HAL SPT SAYA, UANG GAK BISA DITARIK SAMPAI SEKARANG, SAYA AKAN VIRALIN MASALAH INI BIAR TDK ADA KORBAN LAGI..!!
Negarif	Aplikasi apaan ini, mau tarik saldo gak bisa, harus jd akun premium, pas mau verifikasi data ektp daftar akun premium ternyata no ektp salah, lhaa trus apa nya yg salah, d ektp emang itu nmr nya, aduh, gmn ni mau cairin duit di bukalapak saya,, aplikasi kampret
Negarif	Semua sudah bagus! Aku puas pakek DANA. Cuma 1 masalah Top Up tolong di perbarui. Saldo rekening sudah terpotong tapi saldo DANA tidak masuk. Dan komplain harus nunggu 1x24 jam atau bisa 7x24 jam. Gak masuk akal. Top Up adalah langkah awal buat transaksi jangan di persusah!
Negarif	Apa2an ini...??? Di update ga bisa Di PAKAI MAIN DANA SURPRISE nya...!!! Mending yg SPIN kmrn, biar pun dikit tapi bisa di pakai. Lah ini apa...??? AMPAS!!! UPDATE MALAH NGAMPAS. Keterangan cm kami Optimalkan terus...!!! GAK AKAN GUA GANTI SEBELUM NORMAL DAN BISA DI PAKAI...!!!

## 2.2. Text Processing

*Text Processing* dilakukan dengan menggunakan *Tokenize*, *Transform Case*, *Stopwords (Dictionary)*. Teknik *Tokenize* digunakan untuk menghilangkan tanda baca yang ada pada kalimat review. Tanda yang

digunakan dapat mempengaruhi suatu kalimat, sehingga mempengaruhi persepsi anggapan terhadap suatu ulasan review yang digunakan. Penggunaan Transform Case untuk mengubah karakter huruf, karena karakter dari huruf dapat mempengaruhi kalimat. Selanjutnya penggunaan Stopwords (Dictionary) untuk mengumpulkan atau membuat kamus kosa kata yang bersifat negatif dengan fungsi akan menghilangkan kata negatif seperti kampret, tai, anjir dan sebagainya.

### 2.3. Feature Selection

Seleksi fitur yang digunakan yang digunakan untuk mengolah data menggunakan algoritma Naive Bayes dan K-Nearest Neighbors (KNN).

Algoritma Naive Bayes merupakan algoritma klasifikasi berdasarkan probabilitas dalam statistik yang dikemukakan oleh Thomas Bayes yang memprediksi peluang di masa depan berdasarkan peluang di masa sebelumnya (teorema Bayes). Metode ini kemudian dikombinasikan dengan “naive” dimana kondisi antar atribut saling bebas tidak berhubungan satu sama lain (Rachmat C & Lukito, 2016)

Algoritma Naive Bayes mempunyai tahapan sebagai berikut (Muthia, 2017).

$$P(x | C) = P(x_1, x_2, \dots, x_n | C)$$

dengan,

C : class

x : vektor dari nilai atribut n

$P(x_i | C)$  : proporsi dokumen dari class

C yang mengandung nilai atribu  $x_i$

Menghitung probabilitas prior untuk tiap class:

$$P(C) = N_j / N$$

$N_j$  : jumlah dokumen pada suatu class

N : jumlah total dokumen

Menghitung probabilitas posterior menggunakan rumus sebagai berikut.

$$P(C | x) = (P(x | C)P(C)) / (P(x))$$

K-Nearest Neighbors (KNN) termasuk kedalam metode supervised learning yang mempunyai tujuan untuk memperoleh pola baru sedangkan metode unsupervised learning bertujuan untuk mendapatkan pola dalam sebuah kata, dalam hal ini berarti metode supervised learning sudah mempunyai label dan yang dimanfaatkan untuk penelitian adalah mengklasifikasikan data (data testing) berdasarkan data yang sudah ada (data training). Sedangkan untuk metode unsupervised learning yaitu mengelompokkan data menjadi cluster sesuai yang telah ditentukan dengan perintah untuk melabelkan cluster tersebut (Pradnyana & Permana, 2018).

Algoritma k-NN adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data training yang jaraknya paling dekat dengan objek tersebut. K-NN adalah sebuah metode untuk mencari kasus

dengan menghitung kedekatan antara kasus baru dan kasus lama. Algoritma k-NN adalah salah satu metode yang digunakan untuk analisis klasifikasi, namun metode k-NN juga digunakan untuk prediksi. Jarak antara dua titik pada data training dan titik pada data testing dapat didefinisikan dengan rumus Euclidean (Noviana, Susanti, & Susanto, 2019), sebagai berikut.

$$d = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2} \quad (4)$$

Keterangan:

- d : jarak Euclidean
- $x_{2i}$  : nilai pada data *testing* ke  $-i$
- $x_{1i}$  : nilai pada data *training* ke  $-i$
- p : banyaknya atribut

Normalisasi Data

Rumus *Min-Max Normalization* sebagai berikut.

$$X = \frac{X - \min(X)}{\max(X) - \min(X)} \quad (5)$$

dengan,

- $X^*$  : data baru
- X : data lama
- min(X) : nilai minimum dari data per kolom
- max(X) : nilai maksimum dari data per kolom

#### 2.4. Model Accuray

Untuk melakukan pengujian terhadap sistem, dilakukan evaluasi akurasi sistem dalam mengklasifikasikan setimen pada

dataset dengan menggunakan confusion matrix Setelah itu, tahap seleksi fitur dilakukan dengan menggunakan Naive Bayes Classifier dan K-Nearest Neighbors. Model akurasi dari proses seleksi fitur maka akan didapatkan hasil Corifusion Metrik dan ROC/AUC dalam bentuk grafik (Rachmat C & Lukito, 2016).

Tabel 2. Confusion Matrix

	true Negatif	true Positif
pred. Negatif	True Negatif (TN)	False Negatif (FN)
pred. Positif	False Positif (FP)	True Positif (TP)

Mempunyai arti, jumlah kategori negatif yang benar sesuai dengan prediksi negatif  
 False Negatif : jumlah kategori negatif yang benar sesuai dengan class positif  
 False Positif : jumlah data positif yang dikategorikan sebagai class negatif  
 True Positif : jumlah data positif yang benar dikategorikan sebagai class positif  
 $Accuracy = (TN + TP) / (TN + FP + FN + TP)$

$$Recall / True Positive Rate = TP / (FP + TP)$$

$$False Positive Rate = FN / (TN + FN)$$

$$Specificy / True Negative = FP / (FP + TP)$$

$$Precision = TP / (FN + TP)$$

$$F-Measure = 2 * TP / (2 * TP + FP + FN)$$

### 3. Results and Discussion

Penelitian dilakukan dengan aplikasi RapidMiner menggunakan Text Processing pada Table 1 yang terdiri dari Tokenize,

Transform Cases, Stopwords (Dictionary).

Hasil pengolahan berupa:

Tabel 3. Text Processing Menggunakan  
 Tokenize, Transform Cases,  
 Stopwords(Dictionary)

<b>Text Review</b>	Apa2an ini...??? Di update ga bisa Di PAKAI MAIN DANA SURPRISE nya...!!! Mending yg SPIN kmrn, biar pun dikit tapi bisa di pakai. Lah ini apa...??? AMPAS!!! UPDATE MALAH NGAMPAS. Keterangan cm kami Optimalkan terus...!!! GAK AKAN GUA GANTI SEBELUM NORMAL DAN BISA DI PAKAI...!!!
	APLIKASI TAI !!!!! MAU UPGRADE KE PREMIUM UDAH UPLOAD KTP RATUSAN KALI TETEP AJA GAK KEDETECT !!!! NYESEL DAH UDAH DOWNLOAD !! GAK GUNA !!!!
	Aplikasi apaan ini, mau tarik saldo gak bisa, harus jd akun premium, pas mau verifikasi data ektp daftar akun premium ternyata no ektp salah, lhaa trus apa nya yg salah, d ektp emang itu nmr nya, aduh, gmn ni mau cairin duit di bukalapak saya,, aplikasi kampret
<b>Tokenize</b>	Apa an ini Di update ga bisa Di PAKAI MAIN DANA SURPRISE nya Mending yg SPIN kmrn biar pun dikit tapi bisa di pakai Lah ini apa AMPAS UPDATE MALAH NGAMPAS Keterangan cm kami Optimalkan terus GAK AKAN GUA GANTI SEBELUM NORMAL

<b>Transform Cases</b>	DAN BISA DI PAKAI APLIKASI TAI MAU UPGRADE KE PREMIUM UDAH UPLOAD KTP RATUSAN KALI TETEP AJA GAK KEDETECT NYESEL DAH UDAH DOWNLOAD GAK GUNA Aplikasi apaan ini mau tarik saldo gak bisa harus jd akun premium pas mau verifikasi data ektp daftar akun premium ternyata no ektp salah lhaa trus apa nya yg salah d ektp emang itu nmr nya aduh gmn ni mau cairin duit di bukalapak saya aplikasi kampret
	apa an ini di update ga bisa di pakai main dana surprise nya mending yg spin kmrn biar pun dikit tapi bisa di pakai lah ini apa ampas update malah ngampas keterangan cm kami optimalkan terus gak akan gua ganti sebelum normal dan bisa di pakai
	aplikasi tai mau upgrade ke premium udah upload ktp ratusan kali tetep aja gak kedetect nyesel dah udah download gak guna aplikasi apaan ini mau tarik saldo gak bisa harus jd akun premium pas mau verifikasi data ektp daftar akun premium ternyata no ektp salah lhaa trus apa nya yg salah d ektp emang itu nmr nya aduh gmn ni mau cairin duit di bukalapak saya aplikasi kampret
<b>Stopwords (Dictionary)</b>	apa an ini di update ga bisa di pakai main dana surprise nya mending yg spin kmrn biar pun dikit tapi bisa di pakai lah ini apa update malah keterangan cm kami

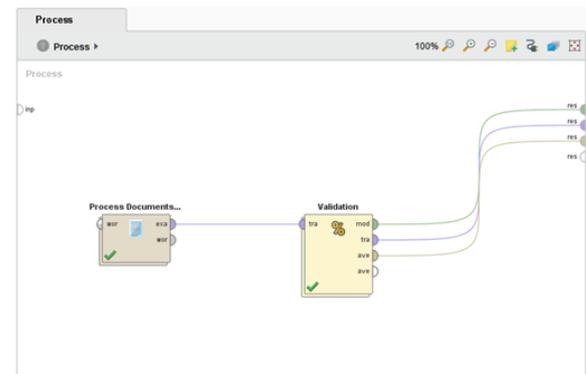
	optimalkan terus gak akan gua ganti sebelum normal dan bisa di pakai
	aplikasi mau upgrade ke premium udah upload ktp ratusan kali tetep aja gak kedetect nyesel dah udah download gak guna
	aplikasi apaan ini mau tarik saldo gak bisa harus jd akun premium pas mau verifikasi data ektp daftar akun premium ternyata no ektp salah lhaa trus apa nya yg salah d ektp emang itu nmr nya aduh gmn ni mau cairin duit di bukalapak saya aplikasi

Tabel 2. menjelaskan mengenai proses teks pada rapidMiner. Pada proses ini menjelaskan text review yang berasal dari reviewer aplikasi Dana. Teks yang mempunyai sifat negatif akan diolah menjadi text yang bersifat positif. Tokenize berfungsi untuk tidak menampilkan tanda baca, karena tanda baca pada text akan berpengaruh text tersebut positif ataupun negatif, tanda baca dapat berupa titik, koma, tanda seru, tanda tanya dan sebagainya. Transform Cases yaitu mengganti karakter huruf yang awalnya huruf besar menjadi huruf kecil, karakter huruf pada text juga berpengaruh terhadap sifat text tersebut, sedangkan Stopwords (Dictionary) mengumpulkan ata membuat kamus kosa kata yang bersifat negatif dengan fungsi akan menghilangkan kata negatif.

Klasifikasi digunakan menentukan apakah sebuah kalimat dari review yang

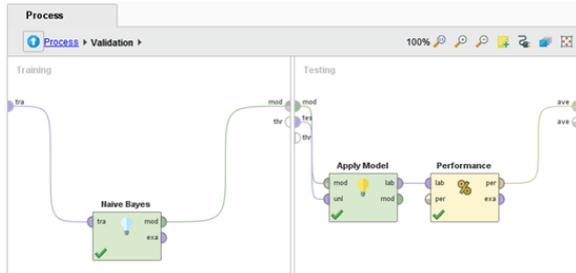
dimasukkan termasuk kelas positif ataupun negatif dengan menggunakan perhitungan yang ada pada algoritma Naive Bayes dan K-Nearest Neighbors. Apabila nilai dari class positif lebih besar dari class negatif maka kalimat tersebut termasuk class positif dan sebaliknya apabila class positif lebih kecil daripada class negatif maka kalimat tersebut tergolong class negatif.

Dalam menentukan hasil penelitian menggunakan algoritma Naive Bayes pada RapidMiner Studio sebagai berikut:



Gambar. 2. Implementasi Process Data dan Validation pada Algoritma Naive Bayes

Fitur yang digunakan berupa *Process Documents* yang digunakan untuk mengolah dokumen dengan memasukkan data review sentimen Positif dan Negatif yang disimpan pada aplikasi Notepad. Di dalam *Process Documents* dimasukkan 3 fitur yaitu *Tokenize*, *Transform Case* dan *Stopwords (Dictionary)*.



Gambar. 3. Implementasi Algoritma Naive Bayes

Algoritma yang digunakan pada pengujian tahap pertama yaitu Algoritma Naive Bayes dengan menggunakan validasi Apply Model dan menggunakan Performance. Pengujian pertama dilakukan menggunakan Algoritma Naive Bayes diperoleh akurasi seperti Tabel IV.

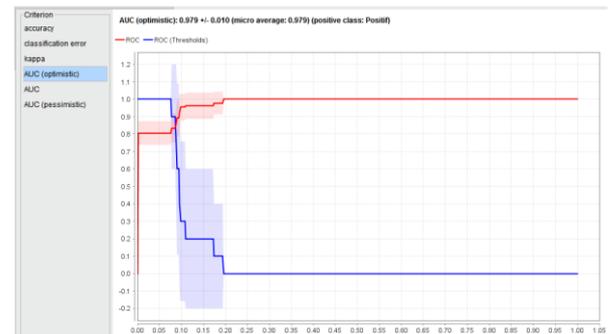
Tabel 4. Akurasi Pengujian Menggunakan Algoritma Naive Bayes

	true Negatif	true Positif
pred. Negatif	385	80
pred. Positif	48	332

Pada tabel 2. untuk review aplikasi Dana menggunakan Algoritma Naive Bayes untuk klasifikasi review negatif sejumlah 385 sesuai dengan prediksi yaitu negatif, untuk klasifikasi review Negatif sejumlah 48 tidak sesuai dengan prediksinya yang memprediksikan positif. Jumlah data review Positif dengan jumlah 80 tetapi ternyata prediksinya Negatif dan 332 data diprediksi review positif sesuai dengan data yang dihasilkan berupa data positif. Proses

klasifikasi menggunakan Algoritma Naive Bayes ini menunjukkan akurasi sebesar 84,76% +/-3,93% dengan rata-rata mikro 84,85%.

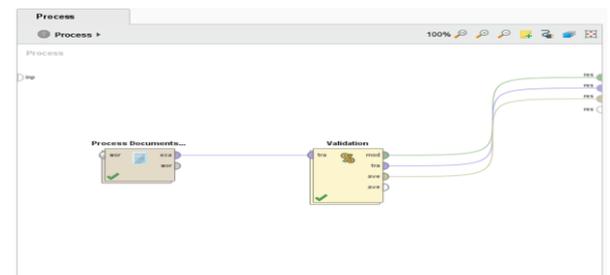
Sedangkan hasil dari ROC/AUC Curve pada Algoritma Naive Bayes sebagai berikut:



Gambar 2. Curve Model Validasi ROC menggunakan Algoritma Naive Bayes

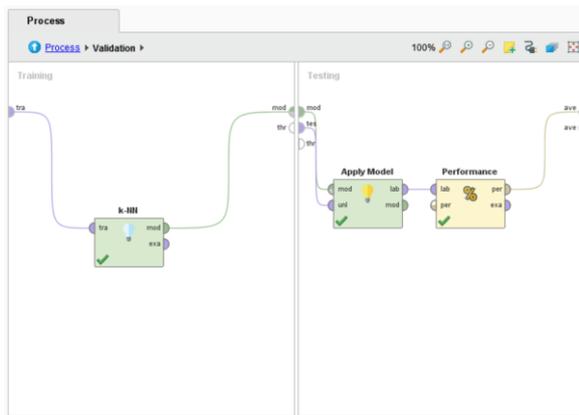
Pada Gambar 2, hasil dari pada penelitian Algoritma Naive Bayes ROC Curve pada penelitian grafik ROC dengan nilai AUC: 0,979 +/- 0,010 (mikro: 0,979) (positive class: Positif), ketentuan ROC menjelaskan apabila nilai ROC mendekati 1 maka nilai semakin baik.

Pengujian kedua dalam klasifikasi review aplikasi Dana menggunakan K-Nearest Neighbors (KNN):



Gambar. 3. Implementasi Process Data dan Validation pada Algoritma K-NN

Fitur yang digunakan berupa *Process Documents* yang digunakan untuk mengolah dokumen dengan memasukkan data review sentimen Positif dan Negatif yang disimpan pada aplikasi Notepad. Di dalam *Process Documents* dimasukkan 3 fitur yaitu *Tokenize*, *Transform Case* dan *Stopwords (Dictionary)*.



Gambar. 4. Implementasi Algoritma KNN

Pengujian pertama dilakukan menggunakan Algoritma K-Nearest Neighbors diperoleh akurasi seperti Tabel 2.

Tabel 2. Akurasi pengujian menggunakan Algoritma K-Nearest Neighbors

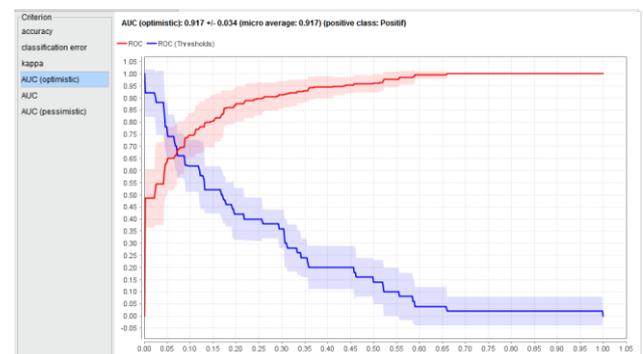
Tabel 5. Akurasi Pengujian Menggunakan Algoritma Naive Bayes

	true Negatif	true Positif
pred. Negatif	364	75
pred. Positif	69	337

Pada Tabel 5. untuk review aplikasi Dana menggunakan K-Nearest Neighbors untuk

klasifikasi review negatif sejumlah 364 sesuai dengan prediksi yaitu negatif, untuk klasifikasi review Negatif sejumlah 69 tidak sesuai dengan prediksinya yang memprediksikan positif. Jumlah data review Positif dengan jumlah 75 tetapi ternyata prediksinya Negatif dan 337 data diprediksi review positif sesuai dengan data yang dihasilkan berupa data positif. Proses klasifikasi menggunakan Algoritma Neural Network ini menunjukkan akurasi sebesar 82,92% +/-4,87% dengan rata-rata mikro 82,96%.

Sedangkan hasil dari ROC/AUC Curve pada Algoritma K-Nearest Neighbors sebagai berikut:



Gambar. 5. Matrix ROC menggunakan Algoritma KNN

Pada Gambar diatas, hasil dari pada penelitian Algoritma K-Nearest Neighbors ROC Curve pada penelitian grafik ROC dengan nilai AUC: 0,917 +/- 0,034 (mikro: 0.917) (positive class: Positif), ketentuan ROC menjelaskan apabila nilai ROC mendekati 1 maka nilai semakin baik.

## **2.2. Kesimpulan**

Pengujian model klasifikasi dengan menggunakan algoritma Naive Bayes dan K-Nearest Neighbors (KNN) yang menggunakan review data analisa sentimen aplikasi Dana berjumlah 232 data, 116 data review positif dan 116 data review negatif. Akurasi yang dihasilkan dengan Algoritma Naive Bayes menghasilkan nilai accuracy sebesar 84,76% +/-3,93% dengan rata-rata mikro 84,85%. Kemudian dilakukan pengujian kembali dengan menggunakan Algoritma K-Nearest Neighbors (KNN) yang menghasilkan akurasi sebesar 82,92% +/-4,87% dengan rata-rata mikro 82,96%. Berdasarkan nilai akurasi tersebut dapat disimpulkan bahwa pengujian analisa sentimen review aplikasi Dana lebih baik menggunakan Algoritma Naive Bayes daripada Algoritma K-Nearest Neighbors. Dalam penelitian ini maka dapat digunakan oleh customer dalam menilai aplikasi dana dan dapat menentukan pilihan dari customer yang telah mereview aplikasi Dana. Penelitian ini juga dapat digunakan untuk pihak manajemen Aplikasi dalam mengelola review pengguna berkaitan dengan sumbang pemikiran untuk mengembangkan Aplikasi.

## **References**

[1] Hanika, I. M. (2015). *FENOMENA PHUBBING DI ERA MILENIA (Ketergantungan Seseorang pada Smartphone terhadap Lingkungannya)*.

4(1), 42–51.  
<https://doi.org/10.14710/interaksi.4.1.42-51>

- [2] Kominfo. (2019). Indonesia Raksasa Teknologi Digital Asia. Retrieved from [https://kominfo.go.id/content/detail/6095/indonesia-raksasa-teknologi-digital-asia/0/sorotan\\_media](https://kominfo.go.id/content/detail/6095/indonesia-raksasa-teknologi-digital-asia/0/sorotan_media)
- [3] Muthia, D. A. (2017). Analisis Sentimen Pada Review Restoran Dengan Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes. *Jurnal Ilmu Pengetahuan Dan Teknologi Komputer*, 2(2), 39–45. <https://doi.org/10.1515/HUMOR.2006.009>
- [4] Noviana, D., Susanti, Y., & Susanto, I. (2019). *Analisis Rekomendasi Penerima Beasiswa Menggunakan Algoritma K-Nearest Neighbor (K-NN) dan Algoritma C4.5*. Retrieved from <http://jurnal.umt.ac.id/index.php/cpu/article/view/1685/1078>
- [5] Pradnyana, G. A., & Permana, A. A. J. (2018). Sistem Pembagian Kelas Kuliah Mahasiswa Dengan Metode K-Means Dan K-Nearest Neighbors Untuk Meningkatkan Kualitas Pembelajaran. *JUTI: Jurnal Ilmiah Teknologi Informasi*, 16(1), 59. <https://doi.org/10.12962/j24068535.v16i1.a696>
- [6] Rachmat C, A., & Lukito, Y. (2016).

- Klasifikasi Sentimen Komentar Politik dari Facebook Page Menggunakan Naive Bayes. *Jurnal Informatika Dan Sistem Informasi Universitas Ciputra*, 2(2), 26–34. <https://doi.org/10.1080/10408398.2013.809690>
- [7] Sugiarti, E. N., Diana, N., & Mawardi, M. C. (2019). E-JRA Vol. 08 No. 01 Februari 2019 Fakultas Ekonomi dan Bisnis Universitas Islam Malang. *E-Jra*, 8(1), 1–13. Retrieved from <http://riset.unisma.ac.id/index.php/jra/article/view/4038/3515>
- [8] Suyanto, S., & Kurniawan, T. A. (2019). Faktor yang Mempengaruhi Tingkat Kepercayaan Penggunaan FinTech pada UMKM Dengan Menggunakan Technology Acceptance Model (TAM). *Akmenika*, 16(1). Retrieved from <https://journal.upy.ac.id/index.php/akmenika/article/view/166>