ESCAF 3rd 2024
p-ISSN : 2962 - 7710
e-ISSN : 3021 - 8594

ANALISIS SENTIMEN APLIKASI DANA PADA GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAIVE BAYES DAN K- NEAREAST NEIGHBOR

Denilia Mertiwi[1*], Budi Santoso[2], Bunga Intan[3]
[1]Program Studi Sistem Informasi, Universitas Bina Insan, Lubuklinggau
e-mail: *[1]2002030042@mhs.univbinainsan.ac.id, [2]budisantoso@univbinainsan.ac.id,
[3]bungaintan@univbinainsan.ac.id

Abstrak

Kemajuan teknologi terutama dalam sektor *Financial Technology* (Fintech), khususnya perkembangan dompet digital di indonesia salah satunya adalah aplikasi DANA. Aplikasi ini memfasilitasi transaksi non-tunai melalui perangkat *mobile*. Analisis sentimen terhadap 2000 ulasan pengguna DANA di *google play store.* Ulasan positif menyoroti kemudahan dan beragamnya fitur, sementara ulasan negatif menyoroti kendala teknis dan pelayanan pelanggan. Analisis sentimen dilakukan menggunakan algoritma *naive bayes* dan *k-nearest neighbor.* Tujuan utama penelitian ini adalah untuk menganalisis sentimen positif dan negatif pada ulasan aplikasi DANA dengan menggunakan algoritma *naive bayes* dan *k-nearest neighbor.* Penelitian ini menggumpulkan 2000 sentimen ulusan pengguna aplikasi DANA dari 2021 hingga agustus 2023 menggunakan library google scraper. Setelah preprocessing, data yang berhasil lolos adalah 1995, dalam klasifikasi dengan 10 *k-fold cross validation* algoritma *naive bayes* mecapai *accuracy* 80,95%, sedangkan *k-nearest neighbor* mencapai akurasi lebih tinggi, yakni 87,07%.

Kata kunci : Aplikasi DANA, Analisis Sentimen, *Naive Bayes, K-Nearest Neighbor*

Abstract

Technological advances, especially in the Financial Technology (Fintech) sector, especially the development of digital wallets in Indonesia, one of which is the DANA application. This application facilitates non-cash transactions through mobile devices. Sentiment analysis of 2000 DANA user reviews on Google Play Store. Positive reviews highlight the ease and variety of features, while negative reviews highlight technical and customer service issues. Sentiment analysis was conducted using naive bayes and k-nearest neighbor algorithms. The main objective of this research is to analyze positive and negative sentiments on DANA app reviews using naive bayes and k-nearest neighbor algorithms. This research collects 2000 sentiments of DANA application user reviews from 2021 to August 2023 using the google scraper library. After preprocessing, the data that passed was 1995, in classification with 10 k-fold cross validation the naive bayes algorithm achieved an accuracy of 80.95%, while the k-nearest neighbor achieved a higher accuracy, namely 87.07%.

Keywords: *DANA Application, Sentiment Analysis, Naive Bayes, K-Nearest Neighbor*

ESCAF 3rd 2024
p-ISSN : 2962 - 7710
e-ISSN : 3021 - 8594

I. PENDAHULUAN

Kemajuan teknologi yang pesat telah membawa dampak luas dengan hadirnya *fintech* (*Financial Technology*). *Fintech* memanfaatkan layanan keuangan untuk memberikan kemudahan dalam melakukan transaksi. Ada banyak fungsi *fintech* yang bisa berkembang dengan cepat. Dompet digital adalah salah satu *fintech* yang sedang berkembang di Indonesia.Di Indonesia, dompet digital atau *e-wallet* sedang mengalami perkembangan pesat. Aplikasi atau layanan ini memungkinkan pengguna untuk membayar barang atau layanan secara *online* dengan menggunakan perangkat *mobile*, seperti *smartphone*(Ii, 2002). Salah satu adanya dompet digital atau *e- wallet* adalah DANA.Aplikasi DANA merupakan aplikasi dompet digital yang dimana pengguna dapat bertransaksi secara non tunai dengan aplikasi ini(Prima, 2021). DANA telah memperoleh popularitas yang besar dari kalangan pengguna dapat dilihat dari *Google Play Store* aplikasi DANA telah didownload sebanyak 100 juta lebih dan memiliki 4 juta ulasan.Pada penelitian ini peneliti mengambil 2000 data ulasan dari *google play store*. Ulasan yang diberikan terdiri dari berbagai macam pendapat positif, negatif maupun netral.

Penelitian terdahulu yang dilakukan Muhammad Farid El Firdaus, dkk. Dengan judul Analisis Sentimen Tokopedia pada Ulasan di *Google Play Store* Menggunakan Algoritma *Naive Bayes Classifier* dan *K-Nearest Neighbor.* Menghasilkan pada nilai *accuracy*, *precision* dan *recall* dari 992 ulasan menghasilkan 403 sentimen negatif dan 589 sentimen positif. Hasil yang diperoleh dari perhitungan performa yang sudah diolah pada metode *naïve bayes* mendapatkan hasil *accuracy* 75,30%, Pada metode *k-nearest neighbor* mendapatkan hasil *accuracy* 86,09%(Firdaus et al., 2022). Ari Putra Wibowo, dkk. Dengan judul Komparasi Metode *Naive Bayes* dan *K- Nearest Nieghbor* Terhadap Analisis Sentimen Pengguna Aplikasi Peduli Lindungi. Menghasilkan Dari hasil eksperimen yang telah dilakukan bisa disimpulkan bahwa, Tingkat akurasi yang lebih baik dihasilkan oleh model klasifikasi dengan algoritma *K-Nearest Neighbor.* Algoritma *K-Nearest Neighbor* menghasilkan tingkat akurasi sebesar 73,33%, sedangkan algoritma *Naive Bayes* tingkat akurasinya sebesar 70,48%(Wibowo et al., 2021). Berdasarkan penelitian-penelitian yang sudah dilakukan sebelumnya penelitian ini akan melakukan analisis sentimen pada ulasan yang terdapat di *google playstore* dengan menggunakan dua metode *Naive Bayes* dan *K- Nearest Neighbor.*

Analisis yang dilakukan pada penelitian ini adalah analisis sentimen Analisis sentimen atau *opinion mining* adalah proses mengolah data tekstual secara otomatis untuk mengetahui apakah opini yang terkandung dalam suatu kalimat adalah positif atau negatif(Khoirul Insan et al., 2023). Penelitian ini menggunakan *text mining*yang bertujuan untuk membandingkan klasifikasi dalam analisis sentimen dari pandangan pengguna aplikasi DANA yang telah ditulis di *playstore* menggunakan metode *Naive Bayes* dan *K-Nearest Neighbor*(Firdaus et al., 2022).

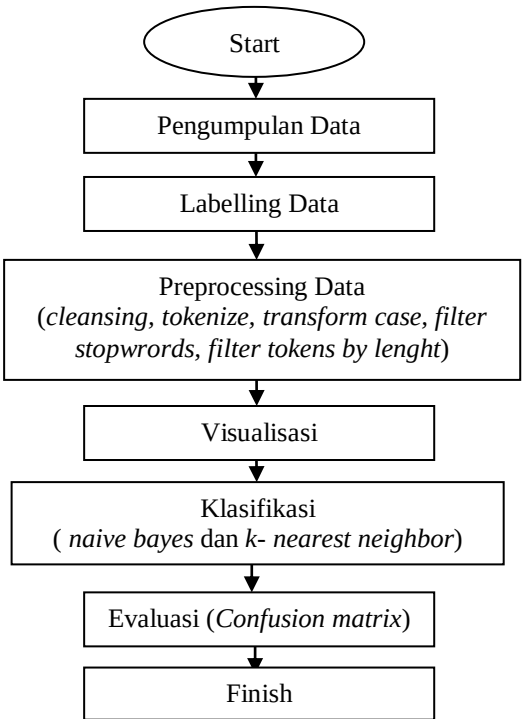
ESCAF 3rd 2024
p-ISSN : 2962 - 7710
e-ISSN : 3021 - 8594

Algoritma *Naïve Bayes* banyak digunakan dalam analisis sentimen karena memiliki keandalan dalam memprediksi kelas dalam dataset pengujian dengan mudah dan cepat. Dalam *Naïve Bayes* berlaku asumsi independensi yakni memiliki performa yang lebih baik dari algoritma lainnya karena memerlukan training data yang lebih sedikit namun berkinerja baik terutama dalam data katagorikal(Indrayuni et al., 2021). Algoritma *K-Nearest Neighbor* (KNN) adalah metode data mining untuk klasifikasi pada kumpulan data yang telah diberi label klasifikasi sebelumnya. Secara lebih spesifik, KNN merupakan metode untuk mengklasifikasikan objek berdasarkan data pembelajaran yang memiliki jarak paling dekat dengan objek diuji(Nugraha, 2022).

DANA telah memperoleh popularitas yang besar dari kalangan pengguna dapat dilihat dari *Google Play Store* aplikasi DANA telah didownload sebanyak 100 juta lebih dan memiliki 4 juta ulasan. Oleh karena itu memahami sentimen pengguna terhadap aplikasi DANA adalah hal yang sangat penting.Pada penelitian ini peneliti mengambil 2000 data ulasan dari *google play store*. Pendapat dari setiap pengguna pada aplikasi DANA di *Google Play Store* dapat berpengaruh pada calon pengguna sebagai bahan pertimbangan penggunaan aplikasi, dan juga peneliti berharap dapat memberikan manfaat yang signifikasi bagi perkembangan aplikasi DANA. Adapun tujuan umum dari adanya penelitian ini yaitu untuk melakukan analisa sentimen positif dan negatif pada ulasan aplikasi DANA pada *Google Play Store* menggunakan algoritma *Naive Bayes* dan *K-Nearest Neighbor*.

II. METODOLOGI PENELITIAN

Metode yang digunakan oleh peneliti ini ditampilkan pada gambar 1. Dimulai dari pengumpulan data, labelling data, preprocessing data, visualisai, klasifikasi algoritma *naive bayes* dan *k- nearest neighbor*, evaluasi *confusion matrix*.



Gambar 1. Alur Penelitian

a. Pengumpulan Data

Pengumpulan data dilakukan menggunakan pemrograman phyton.

b. Labelling Data

labelling data dilakukan dengan cara melihat rating pada komentar di *google play store* pada aplikasi DANA . Untuk melihat hasil rating dari komentar dan ulasan tersebut apakah positif, negatif atau netral.

c. Preprocessing Data

Tahap *preprocessing* adalah mengubah data yang berhasil dikumpulkan menjadi format yang diperlukan. Adapun langakah-langkah tahap *preprocessing* yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Cleansing

proses *cleansing* yang digunakan pada beberapa operator untuk pengahapusan simbol, karakter, angka. Pada penghapusan simbol ini operator yang digunakan yaitu operator replace.

ESCAF 3rd 2024
p-ISSN : 2962 - 7710
e-ISSN : 3021 - 8594

2. Tokenize

Tokenizing Pada tahapan *tokenizing*, setiap kata akan dipisahkan berdasarkan spasi yang ditemukan(Apriani et al., 2019).

3. Transform Case

Fitur *transform cases* kita dapat secara otomatis mengubah semua huruf pada teks menjadi huruf kecil semua atau menjadi huruf kapital semua(H, n.d.). Pada penelitian ini peneliti menggunakan huruf kecil semua

4. Filter Stopwords

Pada tahap *filter stopword* dimana tahapan pada text preprocessing dilakukan penghapusan kata-kata yang tidak tidak memiliki makna namun tidak berpengaruh pada suatu kalimat.

5. Filter tokens (by lenght)

Filter token (by length) pada tahap ini dilakukan penghapusan sejumlah kata atau kata singkatan yang miniamal karakternya sudah ditentukan sebelum digunakan. Pada penelitian ini peneliti menggunakan panjang minimal 4 karakter dan maksimal 25 karakter(Khoirul Insan et al., 2023).

d. Visualisasi

Visualisasi adalah cara untuk menyajikan informasi dan pola data melalui grafik, diagram, atau peta. Ini dapat membantu memahami data dengan lebih cepat dan lebih efektif.

e. Klasifikasi

Algoritma klasifikasi *Naive Bayes* merupakan algoritma yang digunakan untuk mencari nilai probabilitas tertinggi untuk mengklasifikasikan data uji pada kategori yang paling tepat(Nugraha, 2022). Algoritma *K-Nearest Neighbors* adalah sebuah algoritma untuk melakukan klasifikasi pada objek berdasarkan data yang jaraknya paling dekat dengan objek tersebut. Tujuan utama dari algoritma ini yaitu mengklasifikasikan suatu obyek berdasarkan atribut-atribut dan training sample(Mara et al., 2021).

f. Evaluasi

Dalam penelitian ini metode pengujian dilakukan menggunakan library yang terdapat pada metode algoritma Naive Baye dan K-Nearest Neighbor yaitu confusion matrix 4 (empat) istilah yang mempresentasikan hasil dari proses klasifikasi.

1) True Positive (TP) Berarti seberapa banyak data yang aktual kelasnya positif, dan model juga memprediksi positif.
2) False Positive (FN) Berarti seberapa banyak data yang aktual kelasnya negatif, namun model memprediksi positif.
3) True Negative (TN) Berarti seberapa banyak data yang aktual kelasnya negatif, dan model memprediksi negatif.
4) False Negative (FN) Berarti seberapa banyak data yang aktual kelasnya positif, namun model memprediksi negatif.

Akurasi, *precision*, dan*recall* dapat dihitung menggunakan informasi yang diberikan dalam confusion matrix.

1) Akurasi

Akurasi adalah ukuran seberapa tepat model dalam mengklasifikasikan data dengan benar, dan nilai akurasi dapat dihitung dengan membagi jumlah prediksi yang benar dengan jumlah total sampel(Ernianti Hasibuan & Elmo Allistair Heriyanto, 2022). Berikut adalah rumus untuk mencari nilai akurasi :

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN}$$

2) Precision

Precision menggambarkan akurasi antara data yang diminta dengan hasil prediksi yang diberikan oleh model, nilai precision yang besar menunjukkan

ESCAF 3rd 2024
p-ISSN : 2962 - 7710
e-ISSN : 3021 - 8594

rendahnya nilai false positive(Ernianti Hasibuan & Elmo Allistair Heriyanto, 2022).Berikut adalah rumus untuk mencari nilai *precision.*

$$Precision = \frac{TP}{FP + TP}$$

3) Recall

Recall adalah ukuran yang menunjukkan rasio pengamatan positif yang diprediksi dengan benar dibandingkan dengan semua pengamatan yang sebenarnya dalam kelas tersebut. Nilai recall yang tinggi menunjukkan rendahnya false negative, yaitu kemungkinan prediksi negatif yang salah oleh model(Ernianti Hasibuan & Elmo Allistair Heriyanto, 2022). Berikut adalah rumus untuk mencari nilai *recall.*

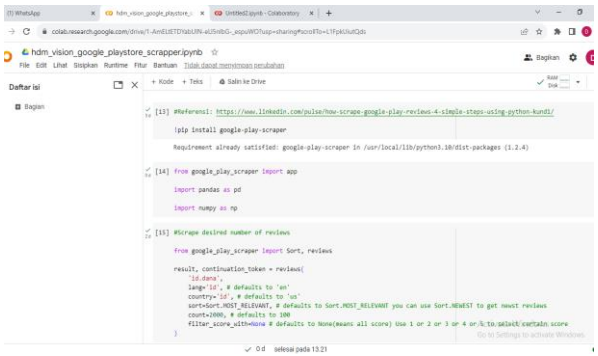
$$Recall = \frac{TP}{FN + TP}$$

III. HASIL DAN PEMBAHASAN

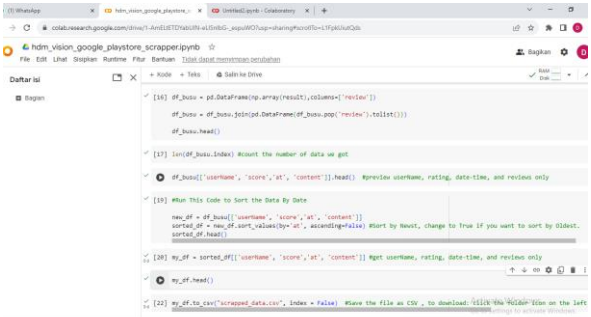
A. Hasil

1. Pengumpulan Data

Pengumpulan data dilakukan menggunakan pemrograman phyton. Langkahpertama dalam pengumpulan data adalah menginstall *library google scraper* di *google colab*. Pada penelitian ini data yang dikumpulkan berjumlah 2000 data. Pengumpulan data dilakukan menggunakan *library* pada *phyton*.



Gambar 2. Source code Pengambilan Data 1



Gambar 3. Source code Pengambilan Data 2

2. Labelling

Setelah mengumpulkan data tahap selanjutnya yaitu melakukan yaitu melakukan pelabelan data secara manual. Dimana proses yang dilakaukan dengan melihat ratting dan membaca satu persatu dari masing-masing ulasan

Tabel 1. Komposisi *Labelling* Data

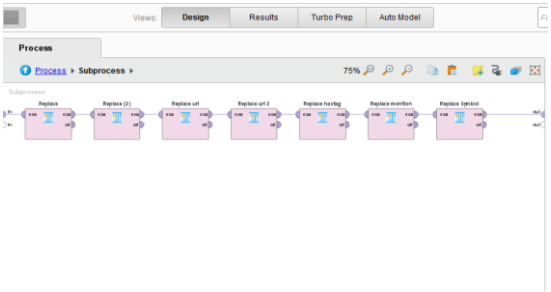
Label	Jumlah
Negatif	1742
Positif	130
Netral	128

3. Preprocessing Data

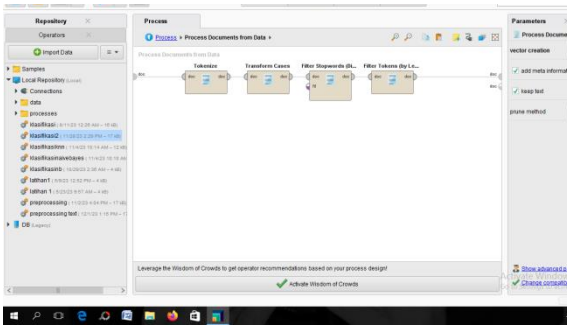
Tahap *preprocessing* adalah mengubah data yang berhasil dikumpulkan menjadi format yang diperlukan. Dengan menggunakan operator yang terdapat di *tools rapid miner.*Adapun langakah-langkah tahap *preprocessing* yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. *Cleansing*
2. *Tokenize*
3. *Transform Case*
4. *Filter Stopsword*
5. *Filter Tokens by Lenght*

ESCAF 3rd 2024
p-ISSN : 2962 - 7710
e-ISSN : 3021 - 8594



Gambar 4. Replace



Gambar 5. Tokenize, Transform Case, filter Stopwords, Filter Tokens By Lenght

Setelah dilakukan proses *preprocessing*, dari 2000 data yang berhasil dikumpulkan menghasilkan 1995 data yang berhasil lolos dari tahap *preprocessing*, hal ini dikarenakan terdapat data duplikat dari sentimen yang telah diberikan oleh pengguna. Berikut data yang lolos dari tahap *preprocessing*. Hasil data yang lolos dari tahap *preprocessing* dapat dilihat di tabel 4.2 dibawah ini.

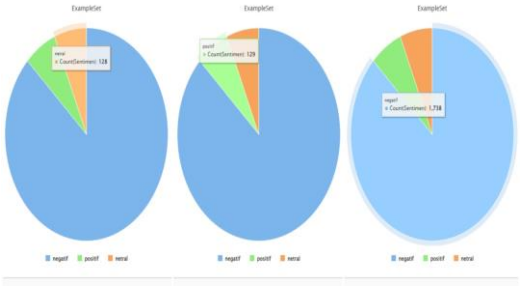
Tabel 2. Komposisi Data yang Berhasil Lolos Tahap *Preprocessing*

Label	Jumlah
Negatif	1738
Positif	129
Netral	128

4. Visualisai

Visualisasi yang digunakan pada penelitian ini adalah visualisi diagram pie dan visualiasi wordcloud.

Berikut gambar di bawah ini merupakan visualisasi diagram pie yang menunjukkan jumlah masing-masing sentimen.



Gambar 5. Diagram Pie Jumlah Sentimen

Gambar 6. Wordcloud Negatif

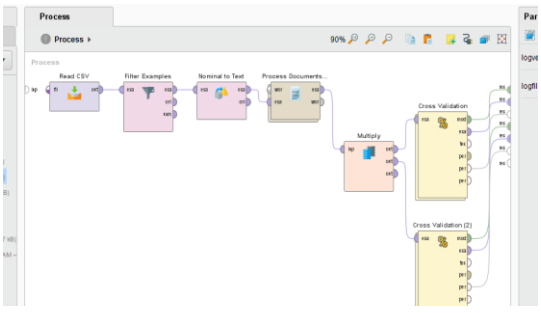
Gambar 7. Wordcloud Positif

Gambar 8. Wordcloud Netral

ESCAF 3rd 2024
p-ISSN : 2962 - 7710
e-ISSN : 3021 - 8594

5. Klasifikasi

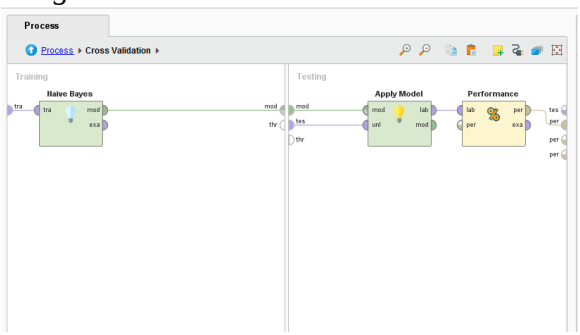
Dalam tahap pembuatan model ini, membuat model dengan kalsifikasi dataset untuk sentimen yang telah melalui *preprocessing*. Dalam tahap penelitian dua algoritma ini didukung oleh *tools rapidminer*. Terdapat beberapa operatordiantaranya "*read csv*" yang berfungsi menampung data, lalu dihubungkan denganoperator "*filter example*" yang berfungsi menyaring atau memfilter baris tertentu dalam dataset, kemudian dihubungkan dengan operator "*Nominal to Text*"untuk mengubah *nominal ke text*, lalu dihubungkan ke "*Process Document From Data"* yang didalamnya terdapat subproses *preprocessing* seperti t*okenize*, *transfrom case*, *filter stopword* dan *filter tokens by lenght*. Kemudian dihubungkan dengan operator *multiply* yang berfungsi untuk melakukan dua sekaligus perhitungan evaluasi algoritma naive bayes dan *k-nearest neighbo*r, lalu dihubungkan ke operator *cross validation*, Pada pengujian ini menggunakan k-f*old cross* v*alidation* berguna untuk menghitung kinerja proses suatu algoritma yang membagi dataset jadi beberapa *k*-buah secara acak serta mengelompokkan data itu sebanyak *k-fold*. Pada penelitian ini memakai metode 10-*fold cross validation* Cara metode 10-*fold cross validation* bekerja yakni membagi data jadi 10-*fold* model yang berukuran sama sehingga memiliki 10 subset data yang masing-masing digunakan 9 subset data latih dan 1 subset untuk data uji dengan 10 kali iterasi (Com, 2022).



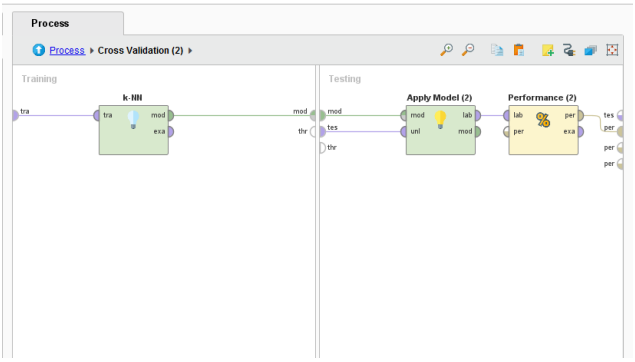
Gambar 9. Model Klasifikasi Algoritma

6. Evaluasi

Tahap ini dilakukan untuk menghitung evaluasi dari model *naive bayes* dan *k-nearest neighbor* dengan *cross validation* yang didalamnya dihubungkan dengan *apply* model dan performance untuk dihitung terhadap hasil akurasi dari algoritma *naive bayes* dan *k-nearest neighbor*.



Gambar 10. Model Klasifikasi Algoritma NB



Gambar 11. Model Klasifikasi Algoritma K-NN

Setelah melalui proses validasi menggunakan algoritma *naive bayes* dan *k-nearest neighbor*, proses evaluasi akan menetukan akurasi untuk masing-masing model.

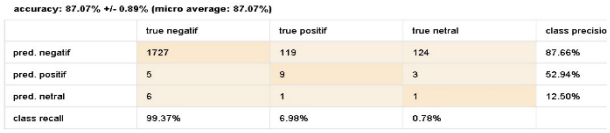
accuracy: 80.95% +/- 1.80% (micro average: 80.95%)

	true negatif	true positif	true netral	class precision
pred. negatif	1591	101	104	88.59%
pred. positif	72	12	12	12.50%
pred. netral	75	16	12	11.65%
class recall	91.54%	9.30%	9.38%	

Gambar 12. Hasil Evaluasi Algoritma NB

ESCAF 3rd 2024
p-ISSN : 2962 - 7710
e-ISSN : 3021 - 8594

Hasil klasifikasi terhadap model algoritma *naive bayes* didapatkan hasil *accuracy* sebesar 80,95%. *Precision* negatif 88.59%, *precision* positif 12,50% dan *percision* netral 11,65%. *Recall*negatif 91,54%, *recall* positif 9.30%, dan *recall* netral 9,38%.



accuracy: 87.07% +/- 0.89% (micro average: 87.07%)

	true negatif	true positif	true netral	class precisio
pred. negatif	1727	119	124	87.66%
pred. positif	5	9	3	52.94%
pred. netral	6	1	1	12.50%
class recall	99.37%	6.98%	0.78%	

Gambar 13. Hasil Evaluasi Algoritma K- NN

Berdasarkan gambar 4.33, bahwa hasil klasifikasi terhadap model algoritma *k-nearest neighbor* didapatkan hasil *accuracy* sebesar 87,07%. *Precision* negatif 87,66%, *precision* positif 52,94% dan *percision* netral 12,50%. *Recall* negatif 99,37%, *recall* positif 6.98%, dan *recall* netral 0,78%.

2. Pembahasan

Masalah yang dibahas pada penelitian ini adalah bagaimana menganalisis ulasan pengguna aplikasi DANA pada *Google Play Store* menggunakan algoritma *Naive Bayes* dan *K-Nearest Neighbor*.

Langkah awal yaitu pengumpaulan data penelitian menggunakan *library google scraper*. Data yang berhasil dikumpulkan dari tahun 2021 sampai dengan bulan agustus 2023. Data penelitian aplikasi DANA yang berhasil didapatkan berjumlah 2000 Sentimen yang telah diberi oleh pengguna aplikasi DANA di google play store. Proses selanjutnya yaitu *labelling* data. *Labelling* data dilakukan secara manual, dengan melihat ratting dan membaca isi sentimen satu persatu apakah sentimen tersebut negatif, positif maupun netral. Hasil dari *labelling* data yaitu untuk sentimen negatif 1742, untuk sentimen positif 130, untuk sentimen netral 128.

Yang kemudian dilakukan *preprocessing* pada data yang terlah berhasil didapatkan dari 2000 data menjadi 1995 data yang berhasil lolos dari tahap *peprocessing* dengan jumlah sentimen negatif sebanyak 1738, sentimen positif sebanyak 129, dan sentimen netral sebanyak 128, hal ini dikarenakan terdapat data duplikat atau sentimen yang sama yang telah diberikan oleh pengguna. Setelah dilakukan*preprocessing*maka proses selanjutnya yaitu klasifikasi dan evaluasi menggunakan metode *naive bayes* dan *k-nearest neighbor*.

Klasifikasi diawali dengam membuat model klasifikasi pada operator-operator yang terdapat di tools rapidminer dimulai dari operator *read csv*, *filter example*, *Nominal to Text*, *Process Documment From Data*, *MultiPly*, *Cross Validation* dengan menggunakan 10 *k-fold*. Kemudian melakuan model evaluasi dengan menerapkan algoritma *naive bayes* dan *k-nearest neighbor*.

Pada hasil klasifikasi algoritma *naive bayes* menghasilkan *accuracy* sebesar 80,95%. *Precision* negatif 88.59%, *precision* positif 12,50% dan *percision* netral 11,65%. *Recall* negatif 91,54%, *recall* positif 9.30%, dan *recall* netral 9,38%. Sedangkan hasil evaluasi algoritma *k- nearest neighbor* mengasilkan *accuracy* sebesar 87,07%. *Precision* negatif 87,66%, *precision* positif 52,94% dan *percision* netral 12,50%. *Recall* negatif 99,37%, *recall* positif 6.98%, dan *recall* netral 0,78%.

Dari perbandingan algoritma antara *naive bayes* dan *k-nearest neighbor* dapat diketahui bahwa algoritma *k-nearest neighbor* lebih tinggi dengan mendapatkan hasil *accuracy* sebesar 87,07%. Sedangkan pada algoritma *naive bayes* mendapatkan hasil *accuracy* sebesar 80,95%.

ESCAF 3rd 2024
p-ISSN : 2962 - 7710
e-ISSN : 3021 - 8594

IV. KESIMPULAN

Adapun kesimpulan dari penelitian yang telah dilakukan mengenai analisis sentimen aplikasi DANA pada *google play store* menggunakan algoritma *naive bayes* dan *k-nearest neighbor* adalah sebagai berikut :

1. Data penelitian diperoleh melalui pengumpulan menggunakan *library google scraper*, dengan total 2000 sentimen dari tahun 2021 hingga agustus 2023. Hasil *labelling* menunjukkan untuk sentimen negatif sebanyak 1742, sedangkan untuk sentimen positif sebanyak 130, dan untuk sentimen netral sebanyak 128.
2. Data yang didapatkan setelah melakukan tahap *preprocessing* dari 2000 data menjadi 1995 data yang berhasil lolos dari tahap *preprocessing* dengan sentimen negatif sebanyak 1738, sedangkan sentimen positif sebanyak 129, dan sentimen netral sebnayak 128.
3. Hasil klasifikasi dengan menggunakan 10 *k-fold cross validation* mengahasilkan klasifikasi algoritma *naive bayes* menghasilkan *accuracy* sebesar 80,95%. *Precision* negatif 88.59%, *precision* positif 12,50% dan *percision* netral 11,65%. *Recall* negatif 91,54%, *recall* positif 9.30%, dan *recall* netral 9,38%. Sedangkan hasil evaluasi algoritma *k- nearest neighbor* mengasilkan *accuracy* sebesar 87,07%. *Precision* negatif 87,66%, *precision* positif 52,94% dan *percision* netral 12,50%. *Recall* negatif 99,37%, *recall* positif 6.98%, dan *recall*netral 0,78%.
4. Dari perbandingan hasil, dapat dilihat bahwa algoritma *k- nearest neighbor* memberikan *accuracy* yang lebih tinggi 87,07% dibandingkan dengan algoritma *naive baye* 80,95% dalam klasifikasi sentimen ulasanpengguna aplikasi DANA di *google play store*. Oleh karena itu dapat disimpulkan *k-nearest neighbor* adalah algoritma terbaik dibandingkan dengan algoritma *naive bayes*

V. SARAN

Untuk meningkatkan hasil pengembangan pada penelitian selanjutnya, ada beberapa saran yang mungkin bisa dilakukan pada peneiti selanjutnya agar mendapatkan hasil yang lebih maksimal yaitu :

1. Memilih model klasifikasi atau metode algoritma yang berbeda untuk melakukan klasifikasi.
2. Memilih topik yang berbeda untuk melakukan penelitian selanjutnya tidak hanya aplikasi DANA.
3. Penelitian selanjutnya disarankan menggunakan *library* bahasa pemograman pada *phtyton*.

VI. DAFTAR PUSTAKA

Apriani, R., Gustian, D., Program, S., Sistem, I., Putra, U. N., Indonesia, S., Raya, J., Kaler, C., 21, N., & Sukabumi, K. (2019). Analisis Sentimen dengan Naïve Bayes Terhadap Komentar Aplikasi Tokopedia. *Jurnal Rekayasa Teknologi Nusa Putra*, *6*(1), 54–62. https://rekayasa.nusaputra.ac.id/article/view/86

Com, T. (2022). *Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia*. *21*(1), 52–61.

Ernianti Hasibuan, & Elmo Allistair Heriyanto. (2022). Analisis Sentimen Pada Ulasan Aplikasi Amazon Shopping Di Google Play Store Menggunakan Naive Bayes Classifier. *Jurnal Teknik Dan Science*, *1*(3), 13–24. https://doi.org/10.56127/jts.v1i3.434

ESCAF 3rd 2024
p-ISSN : 2962 - 7710
e-ISSN : 3021 - 8594

Firdaus, M. F. El, Nurfaizah, N., & Sarmini, S. (2022). Analisis Sentimen Tokopedia Pada Ulasan di Google Playstore Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor. *JURIKOM (Jurnal Riset Komputer)*, *9*(5), 1329. https://doi.org/10.30865/jurikom.v9i5.4774

H, A. T. J. (n.d.). *Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining*. 1–9.

Ii, B. A. B. (2002). *Bab I □*. *lim*(2009), 1–25.

Indrayuni, E., Nurhadi, A., & Kristiyanti, D. A. (2021). Implementasi Algoritma Naive Bayes, Support Vector Machine, dan K-Nearest Neighbors untuk Analisa Sentimen Aplikasi Halodoc. *Faktor Exacta*, *14*(2), 64. https://doi.org/10.30998/faktorexacta.v14i2.9697

Khoirul Insan, M. K., Hayati, U., & Nurdiawan, O. (2023). Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di Google Play Menggunakan Algoritma Naive Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(1), 478–483. https://doi.org/10.36040/jati.v7i1.6373

Mara, A. A. P. T., Sediyono, E., & Purnomo, H. (2021). *Penerapan Algoritma K-Nearest Neighbors Pada Analisis Sentimen Metode Pembelajaran Dalam Jaringan (DARING) Di Universitas Kristen Wira Wacana Sumba*. 24–31.

Nugraha, T. P. A. (2022). *Algoritma K-Nn , Naïve Bayes Dan Svm*.

Prima, E. S. (2021). *Evaluasi User Experience Pada Aplikasi E-WalletDengan Metode User Experience Questionnaire*.

Wibowo, A. P., Darmawan, W., & Amalia, N. (2021). *KOMPARASI METODE NAÏVE BAYES DAN K-NEAREST NEIGHBOR TERHADAP ANALISIS SENTIMEN PENGGUNA APLIKASI PEDULILINDUNGI*.