

## Implementasi Metode K-Nearest Neighbor (K-NN) untuk Analisis Sentimen Kepuasan Pengguna Aplikasi Teknologi Finansial FLIP

Sri Rahayu<sup>1,\*</sup>, Yumarlin MZ<sup>1</sup>, Jemmy Edwin Bororing<sup>1</sup>, Rahmat Hadiyat<sup>1</sup>

<sup>1</sup> Program Studi Informatika, Universitas Janabadra, Indonesia

\* Correspondence: ayu.dj@janabadra.ac.id

**Copyright:** © 2022 by the authors

Received: 5 April 2022 | Revised: 15 April 2022 | Accepted: 23 Mei 2022 | Published: 20 Juni 2022

### Abstrak

Fenomena perkembangan teknologi dapat mengubah sistem dalam sektor apapun menawarkan efisiensi, kemudahan akses, kenyamanan dengan biaya yang lebih hemat termasuk dalam sektor keuangan. Aplikasi Flip merupakan salah satu bentuk teknologi finansial yang menyediakan layanan untuk melakukan transfer antar bank bebas biaya admin. Hingga akhir 2021, aplikasi Flip mendapatkan rating 4.9 di Google Play Store. Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap aplikasi Flip sehingga dapat diketahui apakah ulasan pengguna aplikasi Flip juga menunjukkan nilai positif seperti rating yang diterima. Penelitian ini menggunakan rangkaian proses *text mining* terhadap data ulasan pengguna aplikasi Flip di Google Play Store dan algoritma klasifikasi yang digunakan adalah K-Nearest Neighbor dengan pembobotan TF-IDF. Hasil penelitian menunjukkan 77,67% dari data uji benar terklasifikasi ke dalam kelas ulasan positif dengan nilai *precision* dan *recall* tinggi yaitu 82,67% dan 86,92%. Selain itu, dari hasil penerapan metode dalam mengklasifikasikan data ulasan pengguna Flip dengan perbandingan data latih dan data uji sebesar 80%:20% diperoleh tingkat akurasi klasifikasi dengan algoritma K-Nearest Neighbor sebesar 76,68%. Ulasan pengguna terhadap aplikasi Flip menunjukkan hasil yang positif seperti *rating* yang diterima di Google Play Store dan algoritma K-Nearest Neighbor dengan pembobotan TF-IDF yang digunakan untuk proses analisis sentimen pengguna terhadap aplikasi Flip.

**Kata kunci:** finansial; KNN; sentiment; teknologi; TF-IDF

### Abstract

*The phenomenon of technological development can transform systems in various sectors to provide efficiency and convenience at a lower cost, including the financial sector. Flip is a financial service application that makes it easy to transfer money between banks without administrative fees. By the end of 2021, the Flip will have a 4.9 rating on the Google Play Store. The purpose of this study was to analyze user sentiment towards the Flip app to see if flip user ratings were as positive as the ratings received. This study uses a set of text mining processes on the user rating data of the Flip app on the Google Play Store, using the classification algorithm K-Nearest Neighbor with TF-IDF weighting. The results show that 77.67% of the test data are correctly classified as positive evaluation classes, with high accuracy and recall rates of 82.67% and 86.92%, respectively. In addition, from the results of applying the Flip user rating data classification method, the comparison between training data and test data is 80%:20%, and the classification accuracy using the K-Nearest Neighbor algorithm is 76.68%. User reviews of the Flip app have shown positive results, as well as the ratings obtained in the Google Play Store and the K-Nearest Neighbor algorithm, TF-IDF weighting process used to analyze user sentiment towards the Flip app.*

**Keywords:** financial; KNN; sentiment; technology; TF-IDF



## PENDAHULUAN

Perkembangan teknologi yang pesat saat ini memunculkan fenomena yang dapat mengubah sistem dalam sektor apa saja yang ada saat ini menawarkan efisiensi, kemudahan akses, kenyamanan dengan biaya yang lebih hemat. Fenomena ini disebut Teknologi Finansial dalam sektor keuangan yang menggabungkan fungsi teknologi dan keuangan (Prawirasasra, 2018). Flip merupakan start-up asal Indonesia yang memberikan warna baru dalam sektor Financial Technology dengan menyediakan layanan untuk melakukan transfer antar bank bebas biaya admin melalui aplikasi yang dapat diakses oleh perangkat berbasis Android dan IOS (Cahyani, 2020).

Selama ini sebuah predikat untuk aplikasi terbaik diberikan kepada aplikasi dengan jumlah unduhan dan pemberian bintang paling tinggi di *Play Store* (Erfina et al., 2020). Berdasarkan profil Flip dalam halaman web flip.id bahwa hingga akhir 2021 (Pertiwi, 2022), aplikasi flip telah digunakan oleh lebih dari 6 juta pengguna dan mendapatkan rating 4.9 di Google PlayStore. Data tersebut menunjukkan Flip merupakan *start-up* lokal yang cukup diminati. Akan tetapi, suatu aplikasi selalu memiliki kekurangan dan kelebihan masing-masing yang dapat menimbulkan berbagai respon dari pengguna aplikasi seperti kepuasan dan kekecewaan terhadap aplikasi tersebut (Giovani et al., 2020). Kolom ulasan komentar terhadap sebuah aplikasi di Google Play menjadi salah satu tempat untuk melontarkan kepuasan dan kekecewaan pengguna atau opini terhadap aplikasi tersebut. Hal tersebut dapat dijadikan bahan analisis sentimen terhadap aplikasi Flip untuk mengetahui apakah jumlah unduhan dan rating aplikasi Flip relevan untuk membuktikan bahwa Flip merupakan aplikasi yang baik dalam sektor Financial Technology.

Analisis sentimen atau opinion mining merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini (Buntoro, 2017). Analisis sentimen dapat dimanfaatkan untuk memprediksi, menganalisa suasana publik, suasana hati, dan gambaran perasaan dari para pengguna internet maupun layanan teknologi yang terkandung dalam teks (Amrullah et al., 2020; Basryah et al., 2021; Que et al., 2020), sehingga dari penelitian ini diharapkan dapat diketahui sentimen pengguna terhadap aplikasi Flip.

Pengungkapan emosi seseorang banyak diekspresikan dalam bentuk tulisan termasuk ungkapan kepuasan maupun kekecewaan yang disampaikan oleh pengguna aplikasi di kolom komentar. Oleh karena itu, proses analisis sentimen yang dilakukan pada penelitian ini menggunakan teknik *text mining*. *Text mining* merupakan proses penambangan data (*data mining*) yang mengambil informasi dari data berupa teks seperti email, dokumen teks, file HTML serta data dari ulasan komentar (Sepri, 2020). *Text mining* juga dapat dimanfaatkan pada beberapa bidang yang memiliki banyak data berupa teks seperti manajemen risiko, deteksi penipuan, *business intelligence*, serta analisis media sosial (Sihombing et al., 2021).

K-Nearest Neighbor (K-NN) merupakan salah satu metode yang dapat digunakan untuk menganalisis sentimen terhadap sebuah produk/aplikasi (Ernawati & Wati, 2018). Algoritma K-NN merupakan metode *lazy learning* dimana tidak ada model yang dipelajari dari data testing, jadi hanya belajar dari contoh uji harus diklasifikasikan (Sari, 2020). KNN menghitung kedekatan antara kasus baru (data test) dan kasus lama (data training) yang berdasarkan pencocokan bobot dari sejumlah fitur yang ada. Setelah jarak terhitung, maka jarak yang paling dekat dengan data latih dianggap memiliki kesamaan (Marisa et al., 2021). Algoritma ini mengklasifikasikan data uji yang belum diketahui kategorinya terhadap data latih yang telah diketahui kategorinya dengan menggunakan rumus euclidean dalam dimensi sebesar k yang mengelilingi data uji (Rozi et al., 2021).

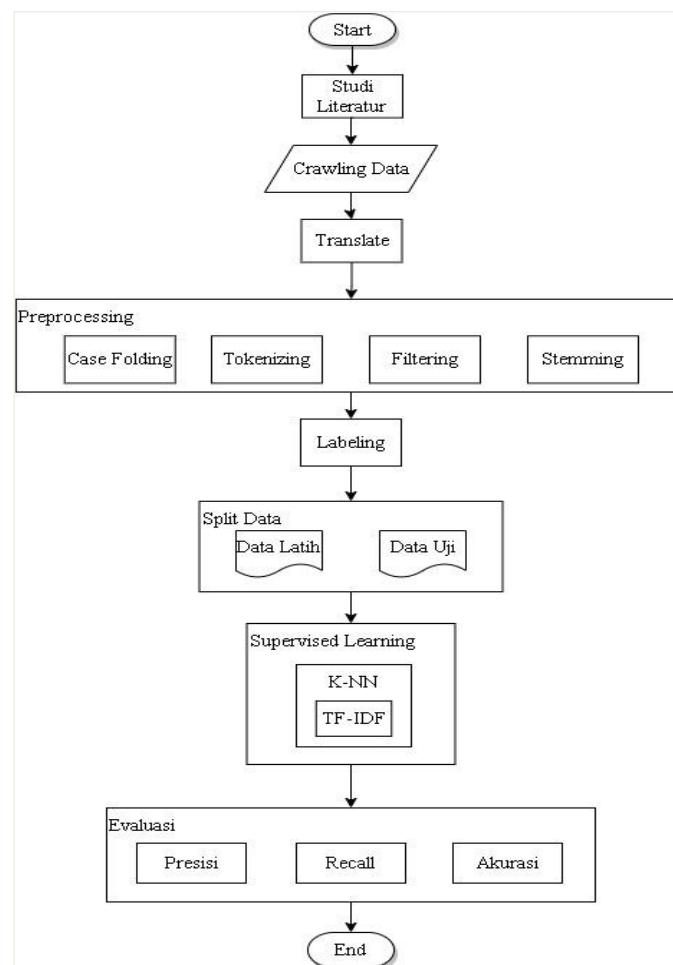
Analisis sentimen ini juga dilakukan pembobotan serta klasifikasi teks yang dapat digunakan untuk menganalisis hubungan antara sebuah kalimat dengan sekumpulan dokumen menggunakan metode *Term Frequency – Inverse Document Frequency* (TF-IDF) (Qaiser &

Ali, 2018). TF-IDF merupakan metode untuk menentukan frekuensi relatif dari setiap kata atau token di mana setiap kata tersebut akan diberikan pembobotan berupa nilai berdasarkan penting atau tidaknya suatu kata dalam dokumen berdasarkan jumlah kemunculan kata dalam suatu dokumen dan mengukur kata-kata tersebut terhadap keseluruhan dokumen yang ada (Rozi et al., 2021).

Aplikasi Flip sebagai start-up dalam bidang teknologi finansial pernah dijadikan sebagai bahan studi kasus untuk membangun *engagement* melalui platform digital (Cahyani, 2020). Adapun beberapa contoh penelitian mengenai analisis sentimen terhadap sebuah produk atau aplikasi yang dijadikan referensi dalam penelitian ini belum ada yang membahas mengenai analisis sentimen pada aplikasi Flip (Basryah et al., 2021; Erfina et al., 2020; Ernawati & Wati, 2018; Giovani et al., 2020; Saputra & Rosiyadi, 2019; Sepri, 2020; Sihombing et al., 2021). Maka dari itu, penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap aplikasi Flip sebagai salah satu produk dalam bidang teknologi finansial.

## METODE

Penelitian ini menggunakan 8 tahapan dalam metode penelitian yaitu (1) *Studi Literatur* (2) *Crawling*, (3) *Translate*, (4) *Preprocessing*, (5) *Labeling*, (6) *Split Data*, (7) *Supervised Learning*, (8) *Evaluasi*. Langkah-langkah penelitian ini dapat dilihat pada Gambar 1.



**Gambar 1.** Alur metode penelitian

Tahap studi literatur dilakukan untuk melihat beberapa referensi baik buku, jurnal penelitian, media sosial, portal berita dan lain sebagainya yang akan digunakan untuk penyusunan penelitian ini. Pada tahap *crawling*, data mentah yang digunakan diambil dari Google Play Store pada laman aplikasi Flip dengan teknik *scraping* (Krotov et al., 2020). Teknik *scraping* pada proses ini dilakukan secara otomatis menggunakan *extention* dari Google Chrome Browser yaitu DataMiner. Selanjutnya pada tahap translate dilakukan perubahan atau mengubah Bahasa Indonesia ke Bahasa Inggris dengan bertujuan untuk memudahkan proses dalam tahapan *preprocessing*.

Pada tahap *Preprocessing* adalah proses perubahan bentuk data yang terstruktur sembarang menjadi data yang terstruktur sesuai kebutuhan untuk proses dalam *text mining* (Takdirillah, 2020). Tahapan *Preprocessing* yang digunakan antara lain: *Case Folding* yaitu tahap mengubah semua huruf dalam dokumen menjadi huruf kecil dan karakter non-huruf akan dihilangkan, *Tokenizing* yang merupakan tahap pemotongan string input berdasarkan tiap kata yang menyusunnya (Afuan, 2018) dan secara garis besar memecah sekumpulan karakter dalam suatu teks ke dalam satuan kata, *Filtering* dengan mengambil kata-kata penting dari hasil tahap *tokenizing* yang dapat dilakukan dengan menghilangkan *stoplist* atau *stopword* (Sugianto & Apandi, 2018), dan *Stemming* yang mentransformasi suatu kata menjadi kata dasarnya (*root word*) dengan menggunakan aturan-aturan tertentu (Senthikumar, Maheswari, 2019).

Sementara itu, di dalam tahapan *Labeling* penulis menggunakan kamus lexicon SentiWordnet untuk menentukan setiap kelas pada setiap komentar. Tahapan ini berjalan dengan cara membuat kamus kata opini (lexicon) terlebih dahulu. Kata-kata yang terdapat pada kamus tersebut akan digunakan untuk mengidentifikasi kata positif dan negatif pada suatu kalimat. Pada tahap ini dilakukan pembagian ke dalam tiga kelas sentimen. Data masuk pada kelas positif jika skor yang didapat  $> 0$ , masuk pada kelas negatif jika skor yang di dapat  $< 0$  dan jika skor = 0 masuk pada kelas netral.

Data hasil *Preprocessing* yang dikategorisasi dan diklasifikasikan ke dalam sentimen positif dan negatif kemudian dibagi menjadi data latih dan data uji pada data yang telah memiliki label kelas dengan perbandingan 80%:20%. Data latih digunakan untuk pembentukan model klasifikasi (Arifiyanti & Wahyuni, 2020). Data yang digunakan sebagai data latih sebanyak 80% dari keseluruhan data yang didapatkan setelah proses Labeling. Selanjutnya, model klasifikasi yang sudah dilatih akan diuji menggunakan algoritma K-Nearest Neighbor (KNN) untuk mengetahui performa klasifikasi. Data uji yang digunakan sebanyak 20% dari keseluruhan data yang didapatkan setelah proses Labeling.

Pada tahapan *supervised learning*, Algoritma K-Nearest Neighbor (KNN) digunakan sebagai metode klasifikasi terhadap data latih dan data uji dengan pembobotan menggunakan metode *Term Frequency – Inverse Document Frequency* (TF-IDF). Metode TF-IDF merupakan metode untuk menentukan frekuensi relatif dari setiap kata-kata atau token-token yang mana setiap kata tersebut akan diberikan pembobotan berupa nilai berdasarkan penting atau tidaknya suatu kata dalam dokumen berdasarkan jumlah kemunculan kata dalam suatu dokumen dan mengukur kata-kata tersebut terhadap keseluruhan dokumen yang ada .

Selanjutnya pada tahap terakhir yaitu proses evaluasi yang untuk mengetahui akurasi dari algoritma yang digunakan terhadap data untuk menganalisa sentimen pengguna terhadap aplikasi *Financial Technology* Flip (Šimundić, 2009). Nilai akurasi menunjukkan rasio dari dataset yang dapat terklasifikasi dengan benar oleh sistem dari keseluruhan dokumen yang ada.

## HASIL DAN PEMBAHASAN

### Hasil

Data yang digunakan dalam penelitian ini merupakan data ulasan atau review yang diberikan oleh pengguna Flip melalui situs Google Play Store. Data yang disajikan dalam tabel

1 diperoleh dengan menggunakan teknik *scraping* berjumlah 11.002 yang merupakan ulasan dari tanggal 4-10 September 2021.

**Tabel 1.** Data ulasan

<b>Nama Pengguna</b>	<b>Tanggal</b>	<b>Ulasan</b>
Nuri Sahin	02 September 2021	Makin tambah cuan dengan cara isi referral ini RGWJ8568 dijamin bakal past dapat, pokoknya pasti cuan kalian.
faris channel	02 September 2021	Sejauh ini puas menggunakan flip. Untuk transfer dana langsung masuk ke rekening tujuan tanpa biaya transfer antar bank. Ketika transfer dari Bank tertentu ada gangguan, langsung diinformasikan sebelum transaksi sehingga bisa mencari alternatif bank yang lain. Barusan beli token listrik juga cepat prosesnya, biaya admin cuma 500. Kelemahan, saldo baru bisa dipakai untuk topup saja dan saldo harus cukup sesuai nominal, tidak bisa pakai saldo sebagian dan kekurangannya via transfer bank
shofia andari	03 September 2021	Alhamdulillah,setelah komplain di playstore dan chat flip, case saya teratasi dan langsung di bantu via WA. Semoga kedepannya flip semakin baik sistemnya agar semakin nyaman dan aman digunakan

Komentar yang sebelumnya berbahasa Indonesia kemudian diubah menggunakan google translate ke dalam bahasa Inggris. Data ulasan yang sudah diubah ke dalam bahasa Inggris dapat dilihat pada tabel 2. Sementara itu sebanyak 11.002 data yang telah didapatkan dari tahapan *crawling* dan telah ditranslasi, selanjutnya melalui beberapa langkah dalam tahapan *preprocessing* dan menghasilkan 9.373 data. Dan dalam tahap *labeling* didapatkan hasil dari komentar positif sebanyak 6.343 data, komentar netral 1.247 data dan komentar negatif sebanyak 1.783 data. Hasil labeling data terdapat dalam tabel 3.

**Tabel 2.** Hasil translasi data ulasan

<b>Ulasan Sebelum di-translate</b>	<b>Ulasan setelah di-translate</b>
makin tambah cuan dengan cara isi referral ini RGWJ8568 dijamin bakal pasti dapat pokoknya pasti cuan kalian.	<i>the more money you make by filling in this referral, rgwj is guaranteed to be sure to get a complete review.</i>
Alhamdulillah, setelah komplain di playstore dan chat flip, case saya teratasi dan langsung di bantu via WA. Semoga kedepannya flip semakin baik sistemnya agar semakin nyaman dan aman digunakan	<i>thank God, after complaining on the playstore and chat, my flip case was resolved and immediately assisted via wa, hopefully in the future the flip will improve the system so that it is more comfortable and safe to use</i>

Hasil pada tabel 3, merupakan teks beberapa kata positif yang terdeteksi pada kamus lexicon sentiwordnet, seperti “good”, “send”, “make”, dan “success”. Untuk kata negatif yang terdeteksi antara lain, “number” dan “fail”. Masing-masing skor kata dalam komentar kemudian dijumlahkan. Jika jumlah skor >0 maka komentar tersebut masuk ke kelas positif, jika <0 maka masuk ke kelas negatif, sedangkan jika jumlah skornya 0 maka dimasukkan ke kelas netral.

Tahapan *split* data ini menentukan data latih dan data uji ke dalam data yang telah memiliki label kelas, dengan perbandingan 80% : 20%. Proses ini mendapatkan hasil pembagian data berupa data latih sebanyak 7499 dan data uji sebanyak 1874. Metode pemberian bobot ini dilakukan dengan cara melakukan pembobotan yang didasarkan pada banyaknya kata yang muncul di dalam dokumen. Hasil penghitungan TF-IDF dapat dilihat pada tabel 4. Komentar pada dokumen kesatu (D1) pada tabel 4 ditemukan kemunculan kata “*account*” sebanyak 2 kali. Banyaknya kemunculan kata tersebut kemudian dihitung dengan langkah metode pembobotan TF-IDF sehingga didapatkan bobot 0,16094 untuk kata “*account*”.

**Tabel 3.** Hasil pelabelan data

Komentar	Skor per kata	Skor	Kelas
<i>good want send transfer fund destin account number deposit flip virtual account send make complic improv system thank success busi</i>	<i>good (0.81) want (-0.07) send (0.03) account (0.19) number (-0.05) flip (-0.02) virtual (0.03) account (0.19) send (0.03) make (0.36) system (0.02) success (0.14)</i>	1,682	Positif
<i>want send money flip transfer flip send money fail seen transfer chat phone activ actual applic correct stupid disappoint</i>	<i>want (-0.07) send (0.03) flip (-0.02) flip (-0.02) send (0.03) fail (-0.29) actual (0.00) correct (-0.02) stupid (-0.12)</i>	-0,468	Negatif
<i>thank applic process pleas thank</i>	0	0	Netral

**Tabel 4.** Hasil penghitungan TF-IDF

Term	Kemunculan Data			TF			TF.IDF		
	D1	D2	D3	D1	D2	D3	D1	D2	D3
<i>account</i>	2	0	0	0,1	0	0	0,16094	0	0
<i>activ</i>	0	1	0	0	0,0526	0	0	0,08471	0
<i>actual</i>	0	1	0	0	0,0526	0	0	0,08471	0
<i>applic</i>	0	1	1	0	0,0526	0,2	0	0,04823	0,18326

Menentukan hasil dari *true positive*, *true negative*, dan *true netral* pada metode K-Nearest Neighbor menggunakan rumus Euclidean dan diperoleh hasil klasifikasi sentimen positif sebanyak 6.343, sentimen negatif sebanyak 1,783 dan sentimen netral sebanyak 1,247. Adapun hasil keluaran *true positive*, *true negative*, dan *true netral* dapat dilihat pada tabel 5. Hasil ini dapat dilihat bahwa dari 1874 data uji, terdapat 1437 data yang terklasifikasi dengan benar, yaitu 1116 kelas positif, 163 kelas negatif, dan 158 data berada pada kelas netral. Dari hasil pada tabel 5 juga ditunjukkan nilai Presisi untuk kelas data dengan ulasan positif terbesar yaitu 82,67%, sedangkan presisi kelas untuk ulasan negatif 65,46% dan ulasan netral sebesar 57,45%. Nilai *Recall* untuk kelas data dengan ulasan positif sebesar 86,92%, sedangkan *recall* kelas untuk ulasan negatif 46,92% dan ulasan netral sebesar 67,23%.

Selanjutnya, dapat dihitung nilai akurasi algoritma K-Nearest Neighbor sebagai metode klasifikasi dengan membandingkan seluruh data yang terklasifikasi dengan benar, dibandingkan dengan seluruh data uji yang menghasilkan prosentase sebesar 76,68%. Hal ini menunjukkan algoritma K-Nearest Neighbor dengan pembobotan TF-IDF layak digunakan

untuk proses analisis sentimen pengguna terhadap aplikasi Flip yang merupakan salah satu bentuk *Financial Technology*.

**Tabel 5.** Hasil evaluasi klasifikasi

	<i>True Positif</i>	<i>True Negatif</i>	<i>True Netral</i>	<i>Class Precision</i>
Pred.Positif	1116	162	72	82.67%
Pred.Negatif	81	163	5	65.46%
Pred.Netral	87	30	158	57.45%
Class recall	86.92%	46.92%	67.23%	

## Pembahasan

Melalui hasil penelitian yang telah dijabarkan pada bagian sebelumnya, data untuk analisis sentimen yang didapatkan dari teknik scrapping terdapat 10.002 data. Setelah dilakukan praproses data, dihasilkan 9373 data yang selanjutnya dibagi menjadi data latih dan data uji dengan perbandingan 80:20 yaitu 7499 data latih dan 1874 data uji. Dari hasil evaluasi klasifikasi yang dilakukan terhadap 1874 data uji, terdapat 1437 data yang terklasifikasi dengan benar dengan 77,66% pada kelas positif, 11,34% termasuk kelas negatif, dan 11% berada pada kelas netral.

K-Nearest Neighbor (K-NN) merupakan salah satu metode yang layak digunakan untuk menganalisis sentimen terhadap sebuah produk/aplikasi. Hal ini dibuktikan dengan tingkat akurasi klasifikasi analisis sentimen aplikasi Flip dengan algoritma KNN sebesar 76,68%. Hasil evaluasi klasifikasi yang ditinjau dari nilai Presisi untuk kelas data dengan ulasan positif terbesar yaitu 82,67%. Hal ini menunjukkan bahwa rasio data yang benar mengandung ulasan positif dibandingkan dengan keseluruhan data yang diprediksi mengandung ulasan positif cukup tinggi. Rasio data yang diprediksi mengandung ulasan positif dibandingkan dengan keseluruhan data yang benar mengandung ulasan positif juga tinggi yang ditunjukkan dengan nilai *Recall* 86,92%. Hasil evaluasi klasifikasi tersebut dipengaruhi dari data dokumen yang diambil, tahapan praproses, dan tahap pembobotan dengan TF-IDF. Data yang diambil adalah data komentar pengguna aplikasi Flip di mana masih sering terdapat kesalahan penulisan (*typo*) seperti kurang huruf dan penggunaan kalimat yang kurang sesuai kaidah bahasa sehingga dapat mempengaruhi proses translasi, pelabelan data, dan pembobotan data.

Sesuai hasil penelitian yang dilakukan oleh (Cahyani, 2020) disebutkan bahwa Flip sangat menyadari komunikasi berperan penting dalam keberlangsungan bisnis. Flip menggunakan platform digital untuk menyampaikan segala sesuatu yang baru tentang produk mereka, sekaligus membentuk pemahaman timbal balik melalui percakapan yang dibangun. Beberapa penelitian mengenai analisis sentimen terhadap sebuah produk atau aplikasi yang dijadikan referensi dalam penelitian ini (Basryah et al., 2021; Erfina et al., 2020; Ernawati & Wati, 2018; Giovanni et al., 2020; Saputra & Rosiyadi, 2019; Sepri, 2020; Sihombing et al., 2021) belum ada yang membahas mengenai analisis sentimen pada aplikasi Flip. Untuk membuktikan baiknya keterikatan antara aplikasi Flip dengan penggunaannya, sentimen analisis ulasan pengguna terhadap aplikasi Flip telah dilakukan dalam penelitian ini.

## SIMPULAN

Berdasarkan hasil pengujian dan pembahasan yang dilakukan, maka dapat diperoleh kesimpulan bahwa ulasan pengguna terhadap aplikasi Flip menunjukkan hasil yang positif seperti rating yang diterima di Google Play Store. Selain itu, algoritma K-Nearest Neighbor dengan pembobotan TF-IDF layak digunakan untuk proses analisis sentimen pengguna terhadap aplikasi Flip yang ditunjukkan oleh tingkat akurasi klasifikasi sebesar 76,68%.

Tahapan praproses data perlu diperbaiki sehingga dapat mengatasi kesalahan klasifikasi akibat penulisan komentar pengguna yang tidak sesuai dengan kaidah bahasa yang baik.

## REFERENSI

- Afuan, L. (2018). Sentimen Analisis Di Twitter (Studi Kasus: Resepsi Pernikahan Putri Presiden Ri Ke-7). *Semnasteknomedia Online*, 6(1), 1-6.
- Amrullah, A. Z., Anas, A. S., & Hidayat, M. A. J. (2020). Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square. *Jurnal BITE*, 2(1), 40–44. <https://doi.org/10.30812/bite.v2i1.804>
- Arifiyanti, A. A., & Wahyuni, E. D. (2020). SMOTE: Metode Penyeimbang Kelas pada Klasifikasi Data Mining. *SCAN - Jurnal Teknologi Informasi dan Komunikasi*, 15(1), 34–39. <https://doi.org/10.33005/scan.v15i1.1850>
- Basryah, E. S., Erfina, A., & Warman, C. (2021). Analisis Sentimen Aplikasi Dompot Digital di Era 4.0 pada Masa Pandemi Covid-19 di Play Store Menggunakan Algoritma Naive Bayes Classifier. *Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra*. 1(1), 189–196.
- Buntoro, G. A. (2017). Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter. *Integer Journal*, 2(1), 32–41. <https://doi.org/10.22487/j26204118.2018.v1.i2.11219>
- Cahyani, I. P. (2020). Membangun Engagement Melalui Platform Digital (Studi Kasus Flip sebagai Start-Up Fintech). *Ekspresi dan Persepsi : Jurnal Ilmu Komunikasi*, 3(2), 76. <https://doi.org/10.33822/jep.v3i2.1668>
- Erfina, A., Basryah, E. S., Saepulrohman, A., & Lestari, D. (2020). Analisis Sentimen Aplikasi Pembelajaran Online di Play Store pada Masa Pandemi Covid-19 Menggunakan Algoritma Support Vector Machine. *Seminar Nasional Informatika (SEMNASIF)*, 1(1), 145-152.
- Ernawati, S., & Wati, R. (2018). Penerapan Algoritma K-Nearest Neighbors Pada Analisis Sentimen Review Agen Travel. *Jurnal Khatulistiwa Informatika*, 6(1), 64–69.
- Giovani, A. P., Ardiansyah, A., Haryanti, T., Kurniawati, L., & Gata, W. (2020). Analisis Sentimen Aplikasi Ruang Guru di Twitter Menggunakan Algoritma Klasifikasi. *Jurnal Teknoinfo*, 14(2), 115. <https://doi.org/10.33365/jti.v14i2.679>
- Krotov, V., Johnson, L., & Silva, L. (2020). Legality and Ethics of Web Scraping. *Communications of the Association for Information Systems*, 47, 539–563. <https://doi.org/10.17705/1CAIS.04724>
- Marisa, F., Maukar, A. L., Farhan, A., Widodo, E. A., Sa'adah, I., & Dasilva, R. T. L. (2021). Pengukuran Tingkat Kematangan Kopi Arabika Menggunakan Algoritma K-Nearest Neighbour. *JIMP : Jurnal Informatika Merdeka Pasuruan*, 6(3), 1-5. <http://dx.doi.org/10.37438/jimp.v6i3.346>
- Prawirasasra, K. P. (2018). Financial technology in Indonesia: Disruptive or collaborative? *Reports on Economics and Finance*, 4(2), 83–90. <https://doi.org/10.12988/ref.2018.818>
- Pertiwi, F. L. I. (2022). Mau transfer antar bank gratis? Pakai Flip! <https://flip.id/landing>
- Qaiser, S., & Ali, R. (2018). Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents. *International Journal of Computer Applications*, 181(1), 25–29. <https://doi.org/10.5120/ijca2018917395>
- Que, V. K. S., Iriani, A., & Purnomo, H. D. (2020). Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 9(2), 162–170. <https://doi.org/10.22146/jnteti.v9i2.102>
- Rozi, F., Sukmana, F., & Adani, M. N. (2021). Pengelompokan Judul Buku dengan Menggunakan Algoritma K-Nearest Neighbor (K-NN) dan Term Frequency – Inverse



- Document Frequency (TF-IDF). *JIMP : Jurnal Informatika Merdeka Pasuruan*, 6(3), 1–5. <http://dx.doi.org/10.37438/jimp.v6i3.346>
- Saputra, S. A., & Rosiyadi, D. (2019). Analisis Sentimen E-Wallet Pada Google Play Menggunakan Algoritma Naive Bayes Berbasis Particle Swarm Optimization. *Jurnal Resti (Rekayasa Sistem dan Teknologi Informasi)*, 3(3), 377-382. <https://doi.org/10.29207/resti.v3i3.1118>
- Sari, R. (2020). Analisis Sentimen pada Review Objek Wisata Dunia Fantasi Menggunakan Algoritma K-Nearest Neighbor (K-NN). *EVOLUSI: Jurnal Sains dan Manajemen*, 8(1). <https://doi.org/10.31294/evolusi.v8i1.7371>
- Senthikumar, Maheswari. (2019). Rule Based Morphological Variation Removable Stemming Algorithm. *International Journal of Recent Technology and Engineering*, 8(4), 1809–1814. <https://doi.org/10.35940/ijrte.C6200.118419>
- Sepri, D. (2020). Penerapan Algoritma Naive Bayes Untuk Analisis Kepuasan Penggunaan Aplikasi Bank. *Journal of Computer System and Informatics (JoSYC)*, 2(1), 135-139.
- Sihombing, L. O., Hannie, H., & Dermawan, B. A. (2021). Sentimen Analisis Customer Review Produk Shopee Indonesia Menggunakan Algoritma Naive Bayes Classifier. *Edumatic: Jurnal Pendidikan Informatika*, 5(2), 233-242. <https://doi.org/10.29408/edumatic.v5i2.4089>
- Šimundić, A.-M. (2009). Measures of Diagnostic Accuracy: Basic Definitions. *EJIFCC*, 19(4), 203–211. PubMed.
- Sugianto, C. A., & Apandi, T. H. (2018). *Pengaruh Tokenisasi Kata N-Grams Spam SMS Menggunakan Support Vector Machine* [Preprint]. INA-Rxiv. <https://doi.org/10.31227/osf.io/vjc7k>
- Takdirillah, R. (2020). Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan. *Edumatic : Jurnal Pendidikan Informatika*, 4(1), 37–46. <https://doi.org/10.29408/edumatic.v4i1.2081>