DATAMATIKA

Tersedia secara daring di:
https://conference.ipb.ac.id/datamatika

SENTIMENT ANALYSIS OF MyBCA APPLICATION REVIEWS USING KNN WITH HYBRID FastText AND TF-IDF FEATURE EXTRACTION

ANALISIS SENTIMEN ULASAN APLIKASI MYBCA MENGGUNAKAN KNN DENGAN EKSTRAKSI FITUR HIBRIDA FASTTEXT DAN TF-IDF

Daffa Kyanfaiza Hidayatullah[1‡], Varel Geo Syah Putra[1], Fatih Avicenna Hizir[1], Muhammad Fauzan Nur Rasendriya[1], Harits Abdurrahman Aufa[1],and Cici Suhaeni[1]

[1]Study Program of Statistics and Data Science, IPB University, Indonesia
[‡]corresponding author: daffakyanfaiza@apps.ipb.ac.id

Abstract

The rapid growth of mobile banking services has increased the number of user reviews on digital platforms such as Google Play Store. These reviews can be utilized to evaluate service quality through sentiment analysis. However, the informal nature of user reviews presents challenges in text classification tasks. This study compares FastText and Hybrid FastText–TF-IDF feature representations, as well as K-Nearest Neighbor (KNN) and Fuzzy K-Nearest Neighbor (FKNN) classification methods for sentiment analysis of myBCA application reviews. The dataset consists of 9,445 user reviews collected from Google Play Store. Model evaluation was conducted using 5-fold cross-validation. The preprocessing stage included case folding, noise removal, normalization of non-standard words, tokenization, and stopword removal. The hybrid representation was constructed through feature concatenation between FastText embeddings and TF-IDF weighting, while Chi-Square feature selection was applied to reduce irrelevant features. The results show that the Hybrid FastText–TF-IDF approach consistently outperformed standalone FastText, with an accuracy improvement of approximately 0.3% across all scenarios In addition, KNN slightly outperformed FKNN with an accuracy difference of less than 0.03%, indicating that the fuzzy membership mechanism contributed minimally to relatively well-separated classes. The best

performance was achieved by the Hybrid FastText–TF-IDF with KNN, obtaining 93.56% accuracy using Cosine distance and K=17. These findings indicate that feature representation quality has a greater impact on classification performance than classifier complexity in sentiment analysis of mobile banking reviews.

Keywords: FastText, KNN, *Mobile Banking*, *Sentiment Analysis*, TF-IDF.

Abstrak

Perkembangan layanan *mobile banking* meningkatkan jumlah ulasan pengguna pada platform digital seperti Google Play Store. Ulasan tersebut dapat dimanfaatkan untuk mengevaluasi kualitas layanan melalui analisis sentimen. Namun, karakteristik ulasan yang bersifat informal menjadi tantangan dalam proses klasifikasi teks. Penelitian ini membandingkan representasi fitur FastText dan Hybrid FastText–TF-IDF serta metode klasifikasi K-Nearest Neighbor (KNN) dan Fuzzy K-Nearest Neighbor (FKNN) pada analisis sentimen ulasan aplikasi myBCA. Dataset yang digunakan terdiri dari 9.445 ulasan pengguna yang diperoleh dari Google Play Store. Evaluasi model dilakukan menggunakan 5-fold cross-validation. Tahap praproses meliputi *case folding*, *noise removal*, normalisasi kata tidak baku, tokenisasi, dan *stopword removal*. Representasi *hybrid* dibentuk melalui *feature concatenation* antara embedding FastText dan pembobotan TF-IDF, sedangkan seleksi fitur *Chi-Square* digunakan untuk mengurangi fitur yang kurang relevan. Hasil penelitian menunjukkan bahwa pendekatan *Hybrid* FastText–TF-IDF secara konsisten meningkatkan performa dibandingkan FastText tunggal, dengan peningkatan akurasi sekitar 0,3% pada seluruh skenario. Selain itu, KNN sedikit lebih unggul dibandingkan FKNN dengan selisih akurasi kurang dari 0,03%, yang menunjukkan bahwa pendekatan fuzzy memberikan kontribusi terbatas pada data dengan kelas yang relatif terpisah. Performa terbaik diperoleh pada kombinasi Hybrid FastText–TF-IDF dengan KNN, dengan akurasi 93,56% menggunakan jarak Cosine dan K=17. Temuan ini menunjukkan bahwa kualitas representasi fitur lebih berpengaruh dibandingkan kompleksitas algoritma klasifikasi dalam analisis sentimen ulasan *mobile banking.*

Kata kunci: Analisis Sentimen, FastText, KNN, *Mobile Banking*, TF-IDF.

1. Pendahuluan

Perkembangan layanan digital telah menjadikan aplikasi *mobile banking* sebagai salah satu sarana utama dalam aktivitas perbankan modern. Kualitas layanan aplikasi *mobile banking* dapat tercermin melalui ulasan pengguna pada laman digital, seperti *Google Play Store.* Oleh karena itu, analisis sentimen terhadap ulasan pengguna dapat dimanfaatkan sebagai pendekatan untuk mengevaluasi kualitas layanan digital perbankan. Akbar dan Pratama (2025) menunjukkan bahwa ulasan pengguna aplikasi *myBCA* dapat digunakan untuk mengidentifikasi persepsi pengguna terhadap layanan *mobile banking*.

Ulasan pengguna aplikasi mobile banking umumnya bersifat tidak terstruktur dan mengandung typo, singkatan, serta variasi penulisan informal. Karakteristik tersebut menyebabkan representasi berbasis kata utuh seperti TF-IDF kurang optimal dalam menangkap hubungan semantik antar kata. Penelitian pada domain *mobile banking* di Indonesia masih didominasi pendekatan TF-IDF dan algoritma klasifikasi konvensional. Misalnya, Bimantara dan Zufria (2024) serta Andrian *et al.* (2022) yang menggunakan pendekatan TF-IDF pada analisis sentimen ulasan aplikasi *mobile banking*. Keterbatasan representasi berbasis kata utuh tersebut mendorong penggunaan pendekatan berbasis *subword* seperti FastText yang mampu merepresentasikan hubungan semantik antar kata dengan lebih baik, terutama pada teks informal. Namun, FastText belum bisa mempertimbangkan tingkat kepentingan kata dalam dokumen secara eksplisit (Choi *et al., 2020*).

Berdasarkan permasalahan tersebut, penelitian ini bertujuan menganalisis pengaruh representasi fitur FastText dan Hybrid FastText–TF-IDF terhadap performa klasifikasi sentimen ulasan aplikasi myBCA menggunakan metode KNN dan Fuzzy KNN. Penelitian ini menggunakan pendekatan berbasis tetangga terdekat (*nearest neighbor-based classification*) karena representasi *embedding* seperti FastText mampu mempertahankan hubungan semantik antar kata melalui representasi vektor *dense*, sehingga relevan digunakan pada pengukuran kemiripan antar dokumen. Selain itu, KNN telah menunjukkan performa yang kompetitif pada berbagai penelitian klasifikasi teks berbasis embedding karena mampu memanfaatkan struktur kedekatan lokal data secara efektif (Cunningham & Delany, 2021). Sementara itu, FKNN mampu menangani kemungkinan *overlap* antar kelas sentimen melalui pendekatan derajat keanggotaan, sehingga relevan untuk karakteristik ulasan pengguna yang cenderung ambigu dan tidak terstruktur.

Secara praktis, hasil penelitian ini diharapkan dapat mendukung pengembangan sistem pemantauan otomatis terhadap ulasan pengguna aplikasi *mobile banking*. Informasi sentimen yang dihasilkan dapat membantu penyedia layanan digital dalam mengidentifikasi keluhan pengguna, seperti kendala *login*, transaksi gagal, maupun performa aplikasi, sehingga dapat digunakan sebagai dasar evaluasi dan peningkatan kualitas layanan *platform*.

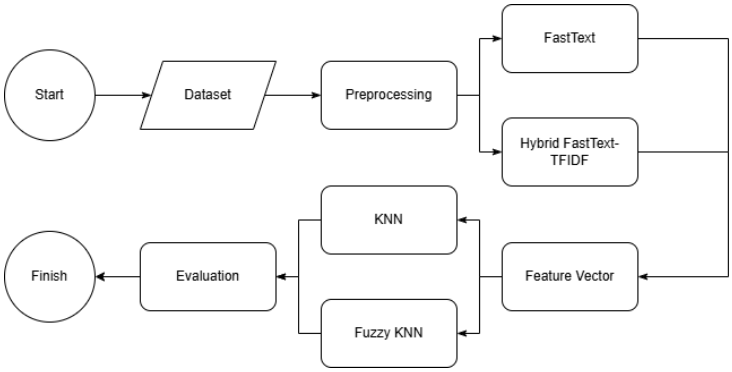
2. Metodologi

2.1 Bahan dan Data

Penelitian ini menggunakan data ulasan pengguna aplikasi myBCA yang diperoleh dari Google Play Store. Data disimpan dalam berkas spreadsheet dengan variabel Komentar (teks ulasan) dan Sentimen (label kelas). Pelabelan dilakukan secara manual oleh lima orang ke dalam dua kategori: sentimen positif dan sentimen negatif. Data yang digunakan sebelum tahap praproses data adalah 9.508 ulasan.

2.2 Metode Penelitian

Penelitian ini mengimplementasikan dua model KNN untuk klasifikasi sentimen, yaitu KNN dan Fuzzy KNN (FKNN), dengan dua jenis representasi fitur: FastText murni dan Hybrid FastText-TF-IDF. Analisis data dilakukan dengan Python melalui laman Kaggle. Alur pemodelan secara keseluruhan ditunjukkan pada Gambar 1.



Gambar 1. Diagram alur penelitian analisis sentimen ulasan aplikasi myBCA.

2.2.1 Praproses Data

Tahap praproses dilakukan untuk membersihkan teks ulasan sebelum ekstraksi fitur (Palomino & Aider, 2022). Proses diawali dengan *case folding* untuk mengubah seluruh teks menjadi huruf kecil, lalu *noise removal* untuk menghapus elemen yang tidak relevan, seperti URL, mention, tanda baca, dan simbol. Kemudian, normalisasi kata tidak baku dengan kamus slang Bahasa Indonesia untuk standarisasi istilah percakapan digital. Setelah itu, teks dipecah menjadi token kata menggunakan pustaka NLTK. Selanjutnya adalah *stopword removal* dengan pustaka PySastrawi untuk menghapus kata umum yang kurang informatif (Andrian *et al*., 2022).

2.2.2 FastText

FastText merepresentasikan informasi *subword* berbasis karakter *n-gram* sehingga model lebih adaptif menangkap hubungan semantik antar kata meskipun terdapat variasi penulisan pada ulasan pengguna aplikasi *mobile banking*. Representasi kata pada FastText dirumuskan pada **Persamaan 1**.

$$v(w) = \left(\frac{1}{|G_w|}\right) \sum_{g}^{G_w} Z_g \quad (1)$$

dengan G_w adalah himpunan n-gram dari kata w dan Z_g adalah vektor representasi n-gram g.

Pada penelitian ini, FastText dilatih menggunakan algoritma *Skip-gram* (sg=1) dengan parameter *epoch* 20 dan *min_count*=2. Selain itu, dilakukan *grid search* terhadap parameter *vector size* {100,200,300,400,500} dan *window size* 1–7 untuk memperoleh konfigurasi model terbaik. Representasi dokumen diperoleh dengan menghitung rata-rata vektor kata pada setiap dokumen, kemudian dilakukan normalisasi menggunakan metode L2.

2.2.3 Hybrid FastText -*Term Frequency-Inverse Document Frequency* (TF-IDF)

TF-IDF mengukur tingkat kepentingan kata berdasarkan frekuensinya dalam dokumen dan kelangkaannya pada keseluruhan korpus (Xiao *et al*., 2024). Metode ini mampu mempertahankan kata yang bersifat diskriminatif terhadap kelas sentimen. Nilai TF-IDF dokumen d dihitung menggunakan **Persamaan 2**.

$$w(t,d) = tf(t,d) \times log\left(\frac{N}{df(t)}\right) \quad (2)$$

dengan *tf(t,d)* adalah frekuensi kata *t* dalam dokumen *d*, *N* adalah jumlah total dokumen, dan *df(t)* adalah jumlah dokumen yang memuat kata *t*.

TF-IDF diekstraksi menggunakan parameter *ngram_range*=(1,2) untuk merepresentasikan unigram dan bigram, serta *min_df*=2 untuk mengurangi pengaruh kata yang jarang muncul. Representasi *hybrid* dibentuk melalui penggabungan fitur FastText dan TF-IDF dengan konatenasi fitur. Namun, karena representasi TF-IDF menghasilkan dimensi fitur yang tinggi, penelitian ini menerapkan seleksi fitur dengan metode *Chi-Square* (χ^2) untuk memilih fitur yang paling relevan terhadap kelas sentimen. Sebanyak 800 fitur TF-IDF dengan nilai χ^2 tertinggi dipilih, dan digabungkan dengan vektor FastText terbaik melalui konkatenasi horizontal sehingga menghasilkan representasi fitur *hybrid*. *Chi-Square* dihitung menggunakan **Persamaan 3**.

$$\chi^2(t,c) = (AD - BC)^2 \times \frac{(A+B+C+D)}{[(A+B)(C+D)(A+C)(B+D)]} \quad (3)$$

2.2.4 K-*Nearest Neighbor* (KNN)

KNN merupakan metode klasifikasi berbasis kemiripan berdasarkan mayoritas kelas dari K tetangga terdekatnya dalam ruang fitur (Halder *et al*., 2024). Pada penelitian, kedekatan antar dokumen dievaluasi menggunakan metrik jarak Euclidean, Manhattan, dan Cosine. Jarak Euclidean dirumuskan sebagai $d^e(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$. Selanjutnya, jarak Manhattan dirumuskan sebagai $d^m(x,y) =$

$\sum_{i=1}^{n}|x_i - y_i|$. Sementara itu, jarak Cosine, dihitung dengan $d^c(x,y) = 1 - \frac{(x \cdot y)}{(||x|| \times ||y||)}$.

2.2.5 Fuzzy K-*Nearest Neighbor* (FKNN)

Fuzzy KNN merupakan pengembangan dari KNN yang memberikan derajat keanggotaan kelas pada setiap data, dan terus disempurnakan pada penelitian terkini (Kumbure & Luukka, 2024). Derajat keanggotaan data uji x dihitung terhadap kelas i dihitung menggunakan **Persamaan 4.**

$$u_i(x) = \sum_{k=1}^{k} \frac{\left[u_i(x_k) \times d(x,x_k)^{-\frac{2}{(m-1)}}\right]}{d(x,x_k)^{-\frac{2}{(m-1)}}} \quad (4)$$

dengan $u_i(x_k)$ menyatakan derajat keanggotaan tetangga ke-*k* terhadap kelas *i*, $d(x,x_k)$ merupakan jarak antara *data uji* dan tetangga ke-*k*, serta *m* adalah parameter *fuzzifier* yang mengendalikan tingkat kekaburan batas kelas. Pada penelitian ini digunakan *m*= 2 sesuai rekomendasi penelitian Halder *et al*.(2024).

2.2.6 Analisis Performa

Dataset yang telah melalui tahap praproses dibagi menjadi data latih dan data uji dengan rasio 80:20 menggunakan *stratified split*. Untuk mengurangi risiko *overfitting*, diterapkan 5-*fold stratified cross-validation* pada data latih. Pada setiap iterasi, empat fold digunakan sebagai data pelatihan dan satu fold digunakan sebagai data validasi secara bergantian. Hasil validasi digunakan untuk menentukan konfigurasi terbaik berdasarkan nilai rata-rata akurasi (Szeghalmy & Fazekas, 2023). Penelitian ini menerapkan tiga skenario eksperimen untuk menganalisis faktor yang memengaruhi performa klasifikasi sentimen yang ditunjukkan di Tabel 1. Pada tahap inisialisasi eksperimen digunakan nilai K=21, metrik jarak Cosine, dan model *baseline* KNN.

Tabel 1: Skenario eksperimen

Skenario	Eksperimen
1	Efek ukuran *Window Size* dan *Vector Size* di ekstraksi fitur
2	Dampak penggunaan jenis matriks jarak pada penentuan nilai K
3	Perbandingan KNN dan Fuzzy KNN berdasarkan Representasi Fitur

Kinerja model dievaluasi dengan beberapa metrik. $Accuracy = \frac{(TP + TN)}{(TP+TN+FP+FN)}$ digunakan karena data yang digunakan memiliki proporsi yang seimbang. Metrik ini mengukur proporsi prediksi yang benar. $Precison = \frac{TP}{TP+FP}$ untuk mengukur ketepatan memprediksi kelas positif (Romadhony *et al.,* 2024). Selain itu, $Recall = \frac{TP}{TP+FN}$ untuk mengukur kemampuan mengenali seluruh data positif (Suryawan *et al*., 2023).

3. Hasil dan Pembahasan

3.1 Praproses Data

Tahap praproses menghasilkan dataset ulasan yang telah terstandarisasi sehingga lebih sesuai untuk tahap ekstraksi fitur dan pemodelan. Hasil akhir menunjukkan bahwa dataset terdiri dari 9.445 ulasan dengan distribusi kelas yang relatif seimbang, yaitu 4.744 sentimen positif dan 4.701 sentimen negatif. Keseimbangan ini penting untuk mengurangi potensi bias model terhadap kelas mayoritas.

Tabel 2: Contoh hasil praproses data

Tahapan	Hasil
Teks awal	“Sangat membantu dalam transaksi sehari hari semoga kedepan lebih responsif lagi:)”
Case Folding	“sangat membantu dalam transaksi sehari hari semoga kedepan lebih responsif lagi:)”
Cleansing	“sangat membantu dalam transaksi sehari hari semoga kedepan lebih responsif lagi”
Stopword Removal	“membantu transaksi sehari hari semoga kedepan responsif”
Stemming	“bantu transaksi hari hari moga depan responsif”
Tokenizing	[“bantu”, “transaksi”, “hari”, “hari”, “moga”, “depan”, “responsif”]

3.2 Skenario 1: Efek ukuran *Window* dan *Vector Size* di Ekstraksi Fitur

Hasil pengujian pada Tabel 3 menunjukkan bahwa kombinasi *window size* 6 dan *vector size* 400 menghasilkan akurasi terbaik sebesar 91,94%. Hasil tersebut menunjukkan bahwa penggunaan konteks kata yang lebih luas membantu FastText menangkap hubungan semantik antar kata pada ulasan pengguna secara lebih baik. Data ulasan *mobile banking* membutuhkan konteks lokal antar kata untuk memahami makna sentimen. Oleh karena itu, peningkatan *window size* memungkinkan model mempelajari dependensi konteks yang lebih baik dan representasi

embedding menjadi lebih informatif dalam membedakan sentimen positif dan negatif.

Tabel 3: Hasil skenario 1

Window	*Vector Size*				
	100	200	300	400	500
1	91,50	91,61	91,61	91,57	91,64
2	91,74	91,75	91,69	91,76	91,72
3	91,73	91,67	91,73	91,68	91,68
4	91,77	91,81	91,84	91,82	91,75
5	91,82	91,71	91,82	91,83	91,76
6	91,82	91,93	91,91	*91,94	91,91
7	91,75	91,66	91,70	91,68	91,67

**Window* size dengan akurasi terbaik

Peningkatan *vector size* tidak selalu menghasilkan performa yang lebih baik. Pada beberapa konfigurasi, peningkatan *vector size* hingga 500 justru menurunkan performa model. Hal ini menunjukkan bahwa representasi berdimensi tinggi belum tentu memberikan informasi tambahan yang signifikan pada dataset ulasan yang relatif pendek. Dimensi vektor yang terlalu besar meningkatkan *sparsity* dan kompleksitas ruang fitur, sehingga model menangkap variasi representasi yang tidak berkaitan dengan sentimen. Akibatnya, proses pengukuran kemiripan pada KNN menjadi kurang stabil. Dengan demikian sesuai dengan penelitian Umer et al. (2023), ukuran konteks yang sesuai berperan penting dalam meningkatkan kualitas embedding dan performa klasifikasi teks. Dalam penelitian ini kombinasi *window size* menengah dan *vector size* yang tidak terlalu besar menghasilkan representasi fitur yang lebih seimbang.

3.3 Skenario 2: Dampak Jenis Matriks Jarak pada Penentuan Nilai K

Berdasarkan konfigurasi optimal pada Skenario 1 (*window size* 6 dan *vector size* 400), eksperimen dilanjutkan untuk mengevaluasi pengaruh metrik jarak dan nilai K pada ruang fitur yang telah dioptimalkan. Hasil pada Tabel 4 menunjukkan bahwa peningkatan nilai K meningkatkan akurasi hingga mencapai titik optimal sebelum akhirnya cenderung stabil. Fenomena ini berkaitan dengan *bias-variance trade-off*, yang disebutkan dalam penelitian Deng *et al*. (2023), yaitu nilai K yang kecil lebih sensitif terhadap *noise*. Sementara itu, nilai K yang lebih besar meningkatkan stabilitas keputusan melalui agregasi tetangga yang lebih luas.

Penelitian Park et al. (2021) menyatakan bahwa jarak Cosine cenderung memberikan performa yang lebih baik pada representasi teks yang bersifat *high-dimensional* dan *sparse*. Jarak Cosine berfokus pada arah distribusi fitur tanpa dipengaruhi panjang vektor, sehingga lebih mampu mempertahankan konsistensi pengukuran kemiripan dokumen pada ruang fitur teks berdimensi tinggi.

Tabel 4: Hasil skenario 2

Nilai K	FastText			Hybrid FastText-TF IDF		
	Euclidean	Manhattan	Cosine	Euclidean	Manhattan	Cosine
1	91,05	91,17	90,31	90,10	91,02	91,41
3	92,47	92,46	91,77	90,17	92,59	92,83
5	92,95	92,95	92,25	89,98	92,92	93,30
7	92,97	93,07	92,33	89,80	93,23	93,33
9	92,98	*93,25	92,32	89,38	93,10	93,39
⋮	⋮	⋮	⋮	⋮	⋮	⋮
17	93,04	93,25	92,43	88,29	93,40	*93,56

* Nilai K dengan Akurasi terbaik

Hasil penelitian ini menunjukkan pola yang serupa. Pada representasi *Hybrid* FastText–TF-IDF, performa tertinggi dicapai oleh jarak Cosine dengan K=17 dan akurasi sebesar 93,56%. Hal tersebut berkaitan dengan karakteristik ruang fitur *hybrid* yang lebih heterogen dan *sparse* akibat pembobotan TF-IDF terhadap dimensi kata berdasarkan tingkat kepentingannya. Kondisi tersebut membuat magnitudo vektor menjadi kurang stabil akibat variasi bobot fitur antar dokumen, sehingga pengukuran berbasis sudut antar vektor lebih representatif dibandingkan jarak absolut.

Sementara itu, pada representasi FastText, performa terbaik diperoleh pada metrik Manhattan dengan K = 9 dan akurasi sebesar 93,25%. FastText menghasilkan representasi embedding yang lebih padat (*dense representation*) sehingga pengukuran berbasis jarak absolut masih mampu merepresentasikan kedekatan antar dokumen secara konsisten. Dengan demikian, efektivitas KNN tidak hanya dipengaruhi oleh nilai K, tetapi juga sangat bergantung pada karakteristik ruang fitur dan metrik jarak yang digunakan.

3.4 Skenario 3: Perbandingan KNN dan Fuzzy KNN berdasarkan Representasi Fitur

Konfigurasi terbaik dari Skenario 2 pada masing-masing representasi fitur digunakan untuk membandingkan performa KNN dan Fuzzy KNN. Hasil pada Tabel 5 menunjukkan bahwa perbedaan kinerja antara KNN dan Fuzzy KNN relatif kecil. Pada FastText, KNN mencapai akurasi 93,25%, sedangkan Fuzzy KNN sebesar 93,26%. Pada representasi Hybrid FastText–TF-IDF, KNN dan Fuzzy KNN masing-masing mencapai 93,56% dan 93,53%, dengan peningkatan hybrid terhadap FastText sebesar 0,33% (KNN) dan 0,29% (Fuzzy KNN). Hasil ini menunjukan bahwa, representasi *hybrid* konsisten memberikan performa terbaik pada kedua metode klasifikasi.

Peningkatan performa pada Hybrid FastText–TF-IDF menunjukkan bahwa integrasi informasi semantik dan pembobotan diskriminatif kata mampu membentuk ruang fitur yang lebih representatif terhadap pola sentimen. FastText efektif menangkap hubungan semantik antar kata melalui representasi *subword* sehingga lebih *robust* terhadap variasi

penulisan informal dan typo, sedangkan TF-IDF menekankan kata-kata yang lebih diskriminatif terhadap kelas sentimen. Kombinasi keduanya menyebabkan dokumen dengan sentimen serupa membentuk *neighborhood* yang lebih stabil dan memperjelas *local decision boundary* antar kelas. Dengan demikian, proses *majority voting* pada KNN menjadi lebih konsisten.

Tabel 5: Hasil skenario 3

Klasifikasi	Ekstraksi Fitur	Konfigurasi Terbaik	Akurasi(%)
KNN	FastText	Manhattan, K = 9	93,25
	Hybrid FastText-TF IDF	Cosine, K=17	93,56
Fuzzy KNN	FastText	Manhattan, K = 9	93,26
	Hybrid FastText-TF IDF	Cosine, K = 11	93,53

Sementara itu, mekanisme derajat keanggotaan pada Fuzzy KNN tidak memberikan kontribusi performa yang substansial dibandingkan KNN klasik. Pada representasi FastText, Fuzzy KNN hanya memberikan peningkatan sangat kecil sebesar sekitar 0,01% dibandingkan KNN. Sementara pada representasi Hybrid FastText–TF-IDF, Fuzzy KNN justru menunjukkan penurunan performa sekitar 0,03%. Hal ini berkaitan dengan karakteristik data yang digunakan. Proses pelabelan dilakukan secara manual sehingga menghasilkan kelas sentimen yang tegas dan tidak ambigu. Label yang jelas (*well-defined classes*) mengurangi potensi *overlapping* antar label sehingga sampel memiliki tetangga terdekat yang homogen. Akibatnya, pendekatan *majority voting* pada KNN sudah cukup optimal dan penambahan derajat keanggotaan fuzzy tidak memberikan kontribusi performa yang substansial.

Selanjutnya dilakukan evaluasi pada model terbaik. Hasil evaluasi pada Tabel 6 menunjukkan bahwa Hybrid FastText–TF-IDF + KNN memiliki performa yang seimbang pada precision (93,62%) dan recall (93,57%). Hal ini menunjukkan bahwa model mampu mengklasifikasikan kedua kelas sentimen secara konsisten dengan tingkat kesalahan (*false positive* dan *false negative*) yang relatif rendah.

Tabel 6: Evaluasi model terbaik

Accuracy	*Precision*	*Recall*
93,56%	93,62%	93,57%

4. Simpulan dan Saran

Penelitian menunjukkan bahwa representasi Hybrid FastText–TF-IDF memberikan performa terbaik pada klasifikasi sentimen ulasan aplikasi myBCA menggunakan metode KNN dan Fuzzy KNN. Berdasarkan tiga skenario eksperimen yang dilakukan, kualitas representasi fitur terbukti memiliki pengaruh yang lebih dominan terhadap performa klasifikasi dibandingkan kompleksitas mekanisme *classifier*. *Hybrid* FastText dan TF-IDF mampu menghasilkan distribusi fitur yang lebih baik dalam memisahkan pola sentimen antar kelas, sehingga meningkatkan konsistensi pengukuran

kemiripan pada KNN. Selain itu, pemilihan nilai K dan metrik jarak juga berpengaruh signifikan terhadap performa model. Sementara itu, perbedaan performa antara KNN dan Fuzzy KNN relatif kecil karena distribusi kelas pada dataset cenderung *well-separated* sehingga *majority voting* pada KNN sudah mampu menghasilkan keputusan klasifikasi yang optimal.

Penelitian selanjutnya disarankan untuk mempertimbangkan fenomena sarkasme dan ironi pada teks ulasan. Hal itu karena aspek tersebut dapat memengaruhi performa model berbasis representasi kata dan embedding. Dataset yang digunakan pada penelitian ini belum secara eksplisit mengakomodasi variasi ekspresi sarkastik. Oleh karena itu, penelitian lanjutan dapat mengeksplorasi proses anotasi maupun penggunaan dataset yang lebih kaya terhadap konteks tersebut.

Daftar Pustaka

Akbar, M. R. M. and Pratama, I. (2025). Sentiment Analysis of MyBCA Application User Reviews using Naive Bayes, Random Forest, and Decision Tree. Sistemasi: Jurnal Sistem Informasi, 14(5), 2464–2478. doi: 10.32520/stmsi.v14i5.5472.

Andrian, B., Simanungkalit, T., Budi, I. and Wicaksono, A. F. (2022). Sentiment Analysis on Customer Satisfaction of Digital Banking in Indonesia. International Journal of Advanced Computer Science and Applications, 13(3). doi: 10.14569/IJACSA.2022.0130356.

Bimantara, M. D. and Zufria, I. (2024). Text Mining Sentiment Analysis on Mobile Banking Application Reviews using TF-IDF Method with Natural Language Processing Approach. JINAV: Journal of Information and Visualization, 5(1), 115–123. doi: 10.35877/454RI.jinav2772.

Cunningham, P. and Delany, S. J. (2021). K-Nearest Neighbour classifiers — a tutorial. *ACM Computing Surveys*, 54(6), Article 128. doi: 10.1145/3459665.

Choi, J., & Lee, S. W. (2020). Improving FastText with inverse document frequency of subwords. *Pattern Recognition Letters*, *133*, 165-172.

Deng, S., Wang, L., Guan, S., Li, M., & Wang, L. (2023). Non-parametric nearest neighbor classification based on global variance difference. *International Journal of Computational Intelligence Systems*, *16*(1), 26.

Halder, R. K., Uddin, M. N., Uddin, M. A., Aryal, S. and Khraisat, A. (2024). Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications. *Journal of Big Data*, 11(1), 113. doi: 10.1186/s40537-024-00973-y.

Kumbure, M. M. and Luukka, P. (2024). Local means-based fuzzy k-nearest neighbor classifier with Minkowski distance and relevance-

complementarity feature weighting. *Granular Computing*. doi: 10.1007/s41066-024-00496-0.

Palomino, M. A. and Aider, F. (2022). Evaluating the effectiveness of text pre-processing in sentiment analysis. Applied Sciences, 12(17), 8765. doi: 10.3390/app12178765.

Park, K., Hong, J. S., & Kim, W. (2021). A methodology combining cosine similarity with classifier for text classification. Applied Artificial Intelligence, 34(5), 396-411.

Romadhony, A., Al Faraby, S., Rismala, R., Wisesti, U. N., & Arifianto, A. (2024). Sentiment Analysis on a Large Indonesian Product Review Dataset. Journal of Information Systems Engineering & Business Intelligence, 10(1).

Suryawan, I. W. B., Utami, N. W., & Fredlina, K. Q. (2023). Analisis Sentimen Review Wisatawan pada Objek Wisata Ubud Menggunakan Algoritma Support Vector Machine. *Jurnal Informatika Teknologi dan Sains (Jinteks)*, *5*(1), 133-140.

Szeghalmy, S., & Fazekas, A. (2023). A comparative study of the use of stratified cross-validation and distribution-balanced stratified cross-validation in imbalanced learning. *Sensors*, *23*(4), 2333.

Umer, M., Imtiaz, Z., Ahmad, M., Nappi, M., Medaglia, C., Choi, G. S., & Mehmood, A. (2023). Impact of convolutional neural network and FastText embedding on text classification. Multimedia Tools and Applications, 82(4), 5569-5585.

Xiao, L., Li, Q., Ma, Q., Shen, J., Yang, Y. and Li, D. (2024). Text classification algorithm of tourist attractions subcategories with modified TF-IDF and Word2Vec. *PLOS ONE*, 19(10), e0305095. doi: 10.1371/journal.pone.0305095.