

PERBANDINGAN ANALISIS SENTIMEN FASILITAS KAMPUS PADA APLIKASI TIKTOK DAN INSTAGRAM FT UMJ

Afrahani Luthfiyah¹, Rully Mujiastuti²

¹Alumni Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jakarta

²Dosen Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah

afrahaniluthfiyah@gmail.com¹, rully@umj.ac.id²

Abstrak

Analisis sentimen digunakan untuk menentukan sentimen atau pendapat positif, negative atau netral yang terkandung dalam teks, termasuk analisis media sosial. Peran media sosial, seperti pada aplikasi *TikTok* dan *Instagram* dalam analisis sentimen sangat penting karena aplikasi ini menjadi tempat di mana orang berbicara secara terbuka tentang berbagai hal, termasuk pada fasilitas di kampus FT UMJ. Pada penelitian ini Penulis mengambil data unggahan *feed dan for your page (fyp)* dari riset terdahulu, dengan *followers Instagram* yang berjumlah 4.758, dan *followers TikTok* yang berjumlah 2.616 pada akun FT UMJ. Tujuannya adalah agar hasil analisis dapat digunakan untuk bahan evaluasi pihak kampus terkait fasilitas kampus dan akademik, sekaligus media sosial yang digunakan untuk publikasi. Penulis menggunakan empat langkah berikut dalam melakukan metodologi penelitian, yaitu; pengumpulan data, *preprocessing data*, analisis sentimen dan klasifikasi data menggunakan metode *Naïve Bayes*. Penulis juga merancang sistem dan tampilan dari aplikasi. Hasil yang diperoleh adalah *F-1 Score*, *Recall* dan *Precision* pada *Instagram* lebih tinggi untuk label negative dibandingkan pada *TikTok*. Nilai akurasi pada *Instagram* bernilai 54% memiliki performa yang lebih baik dari *TikTok* yang memiliki nilai akurasi 33%. Sehingga, *Instagram* dapat direkomendasikan mendapatkan prioritas untuk dikembangkan.

Kata Kunci: *Analisis Sentimen, Instagram, Media Sosial, Naïve bayes, TikTok*

Abstract

Sentiment analysis determines positive, negative, or neutral sentiments or opinions contained in the text, including social media analysis. The role of social media, such as the *TikTok* and *Instagram* applications in sentiment analysis is important because these applications are places where people talk openly about various things, including facilities on the FT UMJ campus. In this study, the author took data from feed uploads and for your page from previous research, with 4,758 *Instagram* followers, and 2,616 *TikTok* followers on the FT UMJ account. The goal is that the results can be used as evaluation material for the campus regarding campus and academic facilities, as well as social media for publication. The author uses the following four steps in conducting the research methodology; data collection, data preprocessing, sentiment analysis, and data classification using the *Naïve Bayes* method. The author also designed the system and appearance of the application. The results are that *F-1 Score*, *Recall*, and *Precision* on *Instagram* are higher for negative labels than on *TikTok*. The accuracy value on *Instagram* is 54%, which has a better performance than *TikTok* which has an accuracy value of 33%. Thus, *Instagram* can be recommended as a priority for development.

Keywords: *entiment Analysis, Instagram, Social Media, Naïve Bayes, TikTok*

1. Pendahuluan Aturan

Era digital menurut (Dwijayanti, 2022) adalah era atau zaman yang mengalami kondisi kemajuan di bidang kehidupan yang serba digital. Hal ini ditandai dengan banyak aplikasi baru yang muncul dan digunakan oleh masyarakat, dengan media sosial sebagai salah satu contoh. Media sosial telah mengubah cara berkomunikasi, berbagi informasi, membangun jejaring, dan mengakses konten. Beragam aplikasi pada media sosial yang digunakan oleh generasi muda dalam kehidupan sehari-harinya, menjadi sebuah gaya hidup baru. Di antaranya adalah *Facebook*, *Twitter*, *Instagram*, *LinkedIn*, *TikTok*, *YouTube*, dan lainnya. (Rasdin, 2021). Hal penting dari fenomena media sosial, adalah adanya interaksi, partisipasi, dan pertukaran konten di dunia maya. (Bulele, 2020)

Dari data (Riyanto, 2023) diketahui bahwa media sosial yang sering digunakan pada saat ini adalah *WhatsApp* - 92,1 %, *Instagram* - 86,5 %, *Facebook* - 83,8 %, *TikTok* - 70,8 %, *Telegram* - 64,3 %, *Twitter* - 60,2 % dan *FB Messenger* - 51,9 % dengan rata-rata usia 16 – 64 tahun. Aplikasi media sosial seperti *TikTok* dan *Instagram* memiliki peran yang signifikan dalam menyampaikan informasi dan pendapat mengenai berbagai aspek kehidupan, termasuk fasilitas kampus, yang terdiri dari fasilitas akademik dan fasilitas layanan dan pendukung.

Komentar yang berada pada publikasi merupakan sebuah tanggapan yang mengandung sentimen. Komentar ini diperoleh dari *feed* dan *for your page (fyp)* pada aplikasi *TikTok* dan *Instagram*, termasuk komentar pada fasilitas. Menurut (Larasati, 2022), fasilitas merupakan suatu yang dapat memudahkan dan memperlancar pelaksanaan suatu usaha, dapat berupa uang, materi, maupun benda fasilitas juga sebagai wadah yang dapat mempermudah.

Penulis mengambil studi kasus Kampus FT-UMJ yang memiliki dua fasilitas yang dikelompokkan; fasilitas akademik dan fasilitas administrasi. Fasilitas akademik berupa sarana dan prasarana yang disediakan kampus untuk membantu proses belajar-mengajar dan memaksimalkan pencapaian akademis mahasiswa. Fasilitas administrasi berupa

sarana yang dipersiapkan untuk mahasiswa agar tidak memiliki kendala dalam mengikuti perkuliahan termasuk dukungan pada mahasiswa yang akan masuk ke dunia kerja.

Melalui data yang diperoleh dari (Luthfiyah, 2023), terdapat *followers Instagram* yang berjumlah 4.758, dan 2.616 *followers* pada aplikasi *TikTok* yang berada di akun FT UMJ. Dari unggahan *feed* dan *for your page (fyp)* tersebut, dilakukan analisis sentimen untuk menghasilkan tanggapan terbaik terkait fasilitas kampus yang berada di aplikasi *TikTok* dan *Instagram*.

Untuk membandingkan sentimen *TikTok* dan *Instagram*, digunakan caption, komentar, dan konten berbasis teks lainnya yang diposting pada aplikasi *TikTok* dan *Instagram* pada akun FT-UMJ. Setelah menganalisis sentimen konten di kedua aplikasi, maka dapat dibandingkan hasilnya untuk melihat apakah ada perbedaan yang signifikan dalam sentimen keseluruhan dari kedua aplikasi tersebut. Untuk dapat menemukan bahwa satu aplikasi memiliki proporsi sentimen positif yang lebih tinggi, sementara aplikasi lainnya memiliki lebih banyak sentimen negatif. Tujuannya adalah hasil analisis sentimen tersebut dapat digunakan untuk bahan evaluasi pihak kampus terkait fasilitas yang ada dan media sosial yang digunakan untuk publikasi.

Metode yang akan digunakan dalam penelitian ini adalah *Naïve Bayes multinomial*. *Naïve Bayes Multinomial* adalah salah satu algoritma yang digunakan dalam klasifikasi data, *Naïve Bayes* merupakan metode yang sering digunakan dalam melakukan analisis sentimen dengan menggunakan teknik prediksi yang berbasis probabilitas berdasarkan teorema Bayes.

Media Sosial

Pengertian media sosial merujuk pada aplikasi-aplikasi online yang memungkinkan pengguna untuk berinteraksi, berbagi konten, dan berkomunikasi dengan pengguna lain melalui berbagai bentuk media seperti teks, gambar, video, dan lainnya. Berikut adalah beberapa definisi media sosial menurut para ahli (Alfiah Zulqornain, 2021)

1. Rheingold (1993): "Media sosial adalah alat komunikasi yang memungkinkan orang untuk berbicara, berbagi informasi, dan

- membuat jejaring secara online."
2. Boyd dan Ellison (2007): "Media sosial adalah situs web yang memungkinkan pengguna untuk membuat profil pribadi, berinteraksi dengan pengguna lain, dan melalui berbagai bentuk media. Situs-situs ini memungkinkan individu untuk mengatur diri dalam jejaring sosial yang lebih besar."

TikTok

TikTok adalah salah satu dari aplikasi media sosial yang memungkinkan pengguna untuk membuat dan berbagi video berdurasi pendek, dengan sebutan *For Your Page (Fyp)*. Biasanya mulai dari 15 detik hingga satu menit dengan *tools like, comment, dan share*. Aplikasi ini populer karena banyaknya berbagai alat kreatif yang dapat digunakan pengguna untuk mengedit dan menyempurnakan video mereka. Pengguna *TikTok* dapat menambahkan musik, filter, efek khusus, dan hamparan teks pada video mereka, dan membagikannya dengan pengikut mereka dan komunitas *TikTok* yang lebih luas. (Bulele, 2020).

Instagram

Menurut (Chaniago, 2020) dalam jurnalnya, seiring berjalannya waktu Dengan berkembangnya teknologi komunikasi, realisasi diri dapat terjadi di dunia maya. Penggunaan media sosial, termasuk *Instagram*. Dengan mengunggah foto atau video, Masyarakat ingin menunjukkan siapa dirinya, sehingga kebutuhannya akan pengakuan sangat besar.

Fasilitas Kampus

Menurut (Larasati, 2022), fasilitas merupakan suatu yang dapat memudahkan dan memperlancar pelaksanaan suatu usaha, dapat berupa uang, materi, maupun benda fasilitas juga sebagai wadah yang dapat mempermudah.

Contoh fasilitas akademik antara lain:

- 1) Ruang perkuliahan,
- 2) Area makan
- 3) Pusat kebugaran
- 4) Laboratorium
- 5) Perpustakaan
- 6) Staff Pengajar

Contoh fasilitas administrasi antara lain:

- 1) Pusat layanan mahasiswa
- 2) PMB (Penerimaan Mahasiswa Baru).

Pre-Processing

Dalam penelitian (Zhafira, 2021) *preprocessing* data digunakan transformasi tata data. Suatu data tidak dapat diproses langsung dengan oleh algoritma, dikarenakan hal tersebut maka dibutuhkan langkah teks *preprocessing* untuk mengubah teks menjadi data numerik dan dapat diolah pada algoritma. Hasil pada *preprocessing* lebih akurat, menggunakan waktu yang singkat, juga memperkecil data tanpa mengubah informasi yang ada pada data (Saifullah, 2017). Proses yang dilakukan dalam *preprocessing* adalah sebagai berikut:

1. Stopword

Dalam jurnal yang ditulis oleh (Elfrida Cristella¹, Yuliant Sibaroni, S.Si. 2020) bahwa *Stopword* adalah proses pembersihan data dari beberapa kata dan karakter yang tidak digunakan, seperti tanda baca. *Stopword* juga sebagai inisialisasi *Stopword* list. Ketentuan dalam melakukan *Stopword* dalam tahapan *case folding* adalah pengecekan ulang kembali, apabila data termasuk kedalam *Stopword*, maka akan diproses kembali ke tahapan *Stemming*, begitupula sebaliknya, apabila data tidak termasuk kedalam *Stopword* maka data tersebut akan diproses ketahap selanjutnya.

2. Stemming

Stemming adalah sebuah proses mengubah kata imbuhan menjadi kata dasar, dengan contoh, lari, berlari, melarikan. Proses *Stemming* menggunakan library *python sastrawi*. Tahap ini merupakan proses yang digunakan untuk mendapatkan bentuk kata dasar dengan cara memisahkan kata dasar dengan imbuhan baik awalan (*prefix*), akhiran (*surfix*), sisipan (*Infixes*) maupun kombinasi awalan dengan akhiran (*Confixes*).

3. Remove Punctuation

Remove punctuation dapat disebut juga sebagai *clean text*. Data yang dihasilkan merupakan data-data yang telah dibersihkan dari karakter-karakter seperti tanda baca (.) (,) (/) dan angka (0,1,2,3,4,5,6,7,8,9).

4. Case folding

Penelitian yang dilakukan (Elfrida Cristella, 2020) mengatakan bahwa Proses *case folding* sebagai proses perubahan huruf

besar menjadi huruf kecil, dan menghilangkan keseluruhan tanda baca pada kalimat. Proses *case folding* dilakukan untuk menyamakan seluruh data agar tidak terjadi kesalahan pada proses tokenizing (Hermanto, 2020).

5. Tokenization

Proses tokenisasi dinamakan karena hasil potongan kalimat disebut sebagai token. Pada tahapan *preprocessing* dilakukan proses pemotongan kalimat dalam sebuah teks menjadi beberapa potongan kata atau karakter yang sesuai dengan keperluan teks. Pemecahan kata dilakukan berdasarkan spasi pada kalimat tersebut, dan mengubah tanda baca menjadi huruf kecil (Sulistio, 2022).

6. TF IDF

Pembobotan kata atau dapat disebut *Term Frequency Inverse Document Frequency (TF-IDF)*, adalah proses pembobotan kata dari hasil seluruh proses preprosesing dengan algoritma untuk melihat berapa banyak kata yang muncul. TF IDF mempunyai arti yang berbeda, TF adalah bobot yang mengacu pada frekuensi kemunculan kata dalam Dokumen, semakin sering Dokumen muncul, maka semakin tinggi skor TF. Begitu pula sebaliknya, IDF yang berartikan seberapa jarang kata dalam Dokumen itu muncul maka akan semakin tinggi skor IDF. (Zhafira, 2021). Oleh karena itu diperlukan perhitungan TFIDF, agar skor dapat diperoleh dengan menggunakan persamaan (Yulita, 2021)

$$Tf - Idf = Tf \cdot \log n/df \quad (1)$$

Dimana:

$Tf - Idf$: bobot suatu dokumen dokumen yang mengandung term tersebut

Tf : frekuensi term pada suatu dokumen

N : jumlah seluruh dokumen

df : jumlah dokumen yang mengandung term tersebut

Adapun rumus yang digunakan untuk menghitung bobot masing-masing dokumen adalah sebagai berikut : (Sasmita & Falani, 2018)

$$Wdt = TF dt * IDF t \quad (2)$$

Dimana:

d : kalimat ke- d

t : kata ke- t dari kata kunci

W : bobot kalimat ke d terhadap term ke- t

Tf : banyaknya kata yang dicari pada sebuah dokumen

IDF : *Inverse Document Frequency*

D : total dokumen

Df : banyak dokumen yang mengandung kata yang dicari

Analisis Sentimen

Metode yang sering digunakan dalam proses ekstraksi dan klasifikasi opini. Analisis suasana hati dilakukan Pengetahuan tentang karakteristik dokumen, kalimat atau entitas / aspek positif, negatif atau netral mengklasifikasikan polaritas teks yang dikandungnya diadili berbagai algoritma yang dapat diimplementasikan dengan melakukan sentimen analisis (Dharmendra, Saputra, and Pramaita, 2019)

Naïve Bayes Multinomial

Naive Bayes Multinomial adalah salah satu algoritma yang digunakan dalam klasifikasi data, *Naïve Bayes* merupakan metode yang sering digunakan dalam melakukan analisis sentimen dengan menggunakan teknik prediksi yang berbasis probalistik berdasar teorema *Bayes*.

Persamaan Metode *Naive Bayes* secara umum:

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (3)$$

Di mana:

H : Prediksi

X : Aktual

$P(H)$: Peluang Prediksi

$P(X)$: Peluang Aktual

$P(X|H)$: Peluang kelas Aktual untuk kelas Prediksi

Berikut adalah rumus NBC dari pertanyaan diatas :

Persamaan Kedua Metode *Naïve Bayes*:

$$P(Hn |Xn) = \frac{P(Xn |Hn)P(Hn)}{P(Xn)} \quad (4)$$

Di mana:

H = Prediksi, X = Aktual, Hn = Kelas Prediksi Ke- n

Dimana:

$H1$ = Positif, $H2$ = Negatif, $H3$ = Netral

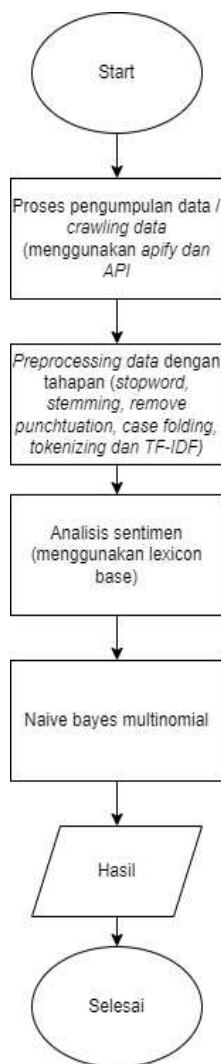
Xn : Kelas Aktual Ke- n

Dimana :

$X1$ = Positif, $X2$ = Negatif, $X3$ = Netral
 $P(X_n)$ = Peluang kelas aktual sampai Ke-n
 $P(H_n)$ = Peluang Kelas Prediksi
sampai Ke-n $P(X_n | H_n)$ = Peluang kelas
Aktual sampai Ke-n untuk kelas Prediksi
sampai Ke-n.

1. Metodologi Penelitian

Penulis menggunakan empat langkah berikut dalam melakukan metodologi penelitian, yaitu; pengumpulan data, *preprocessing data*, analisis sentimen dan klasifikasi data menggunakan metode *Naïve Bayes*. Metodologi penelitian dapat dilihat pada Gambar 1. Berikut.



a. Pengumpulan Data

Gambar 1 Flowchart metode penelitian

Penulis melakukan pengumpulan data melalui *crawling data* yang bertujuan untuk pengambilan data secara otomatis.

Metode *crawling data* melalui *Apify* dan API menjadi teknik yang digunakan untuk mengumpulkan informasi berupa komentar pada *feed* dan FYP dalam aplikasi *TikTok* dan *Instagram* pada akun FT-UMJ. Langkah awal untuk melakukan *crawling data* pada data *Instagram* adalah memasukan komentar *Instagram* dalam URL, proses selanjutnya inisialisasi data *apify client* pada API *Instagram*, setelah API memproses data maka data *crawling* telah didapatkan.

Crawling data pada TikTok

Adapun langkah yang dilakukan untuk *crawling data* pada akun *TikTok* FT-UMJ adalah seperti gambar 2 berikut :



Gambar 2 Crawling data TikTok

Dari hasil *crawling data TikTok* tersebut diperoleh hasil berikut :

Tabel 1 Hasil Crawling Data Pada TikTok

Dok 1	Gym nya kurang besar min
Dok2	kelas yang masih pake bangku kayu tolong diganti min, sama ac di beberapa ruangan ada yg ga dingin. semoga bisa membuat lebih nyaman lagi dlm belajar.
Dok3	Mohon dicek kamar mandi perempuan di belakang masjid ya min, suka ga ada airnya
Dok4	alat di ruang gymnya bagus, cocok untuk latihan pemula
Dok5	perpustakaan bagus banget min udah ganti AC jadi lebih dingin

Crawling Data Instagram

Adapun langkah yang dilakukan untuk *crawling data* pada akun *Instagram* FT-UMJ adalah seperti Gambar 3 berikut:



Gambar 3 Crawling Data Instagram

Dari hasil *crawling data Instagram* tersebut diperoleh hasil Tabel 2 berikut :

Tabel 2 Hasil *Crawling* Data Pada *Instagram*

Dok 1	Gym nya kurang besar min
Dok2	lahan parkir nya luas gak min?
Dok3	Selain fasilitas tempat ada fasilitas lain yg bisa mewadahi mahasiswa nya ga min?
Dok4	unit kesehatan nya punya fasilitas lengkap min?
Dok5	Minimnya fasilitas dalam ruangan kelas
Dok 6	Ada ketentuan buat peminjaman ruang podcast ga ya, kemudian peminjaman ruangan sudah termasuk peminjaman alat alat yang ada di dalam ruangan kah?
Dok 8	perpustakaan nya bagus tapi saran saya lebih bagus lagi untuk di filter ulang buku2 yang sudah lama di ganti dengan buku2 keluaran baru..
Dok 9	Uks nya belum ada perawatnya min
Dok 10	perpusnya bagus👍
Dok 11	Almamater ❤️
Dok 12	Mini gym buka hari apa aja min?
Dok 13	Di ruang podcast bahas apa aja min?
Dok 14	Mini market nya ada yg jaga ga ya ka? Soalnya mau ke sana tapi tutup terus
Dok 15	Bahas fasilitas yang ada di Prodi Teknik Mesin dong min.

b. Preprocessing

Tahap pengolahan selanjutnya melalui *preprocessing*, pengolahan data yang diambil dari data sebelumnya. Tahapan-*preprocessing* sebagai berikut; 1) *Stopword*, 2) *Stemming*, 3) *Remove Punctuation*, 4) *Case folding*, 5) *Tokenizing* dan TF IDF.

Stopword

Pembersihan data dari beberapa kata dan karakter yang tidak digunakan, seperti tanda baca. contoh titik (.), koma (,), titik dua (:), titik koma (;), tanda tanya (?), tanda seru (!), dan tanda petik ("), l. Contoh dari *stopword* adalah “yang”, “untuk”, “di”, “dan”, “atau”. Adapun hasil tampilan dari proses *stopword Instagram* dan *TikTok* pada tabel 3 dan 4 berikut.

Tabel 3 Hasil data *stopword Instagram*

Dok 1	Gym nya kurang besar min
Dok2	lahan parkir nya luas gak min

Dok3	Selain fasilitas tempat ada fasilitas lain yg bisa mewadahi mahasiswa nya ga min
Dok4	unit kesehatan nya punya fasilitas lengkap min
Dok5	Minim nya fasilitas dalam ruangan kelas
Dok 7	Ada ketentuan buat peminjaman ruang podcast ga ya kemudian peminjaman ruangan sudah termasuk peminjaman alat alat ada dalam ruangan kah
Dok 8	perpustakaan nya bagus tapi saran saya lebih bagus lagi untuk filter ulang buku2 sudah lama ganti dengan buku2 keluaran baru
Dok 9	Uks nya belum ada perawat nya min
Dok 10	perpusnya bagus👍
Dok 11	Almamater ❤️
Dok 12	Mini gym buka hari apa aja min
Dok 13	ruang podcast bahas apa aja min
Dok 14	Mini market nya ada yg jaga ga ya ka Soalnya mau ke sana tapi tutup terus
Dok 15	Bahas fasilitas ada ProTeknik Mesin dong min

Tabel 1 Hasil *stopword TikTok*

Dok 1	Gym nya kurang besar min
Dok2	kelas masih pake bangku kayu tolong ganti min sama ac beberapa ruangan ada yg ga nginsemoga bisa membuat lebih nyaman lagi dlm belajar
Dok3	Mohon cek kamar mandi perempuan belakang masjid ya min suka ga ada airnya
Dok4	alat ruang gymnya bagus cocok untuk latihan pemula
Dok5	perpustakaan bagus banget min udah ganti AC aja lebih dingin

Stemming

Tahap ini merupakan proses yang digunakan untuk mendapatkan bentuk kata dasar dengan cara memisahkan kata dasar dengan imbuhan baik awalan (*prefix*), akhiran (*surfix*), sisipan (*Infixes*) maupun kombinasi awalan dengan akhiran (*Confixes*).

Adapun *script* dari *stemming* adalah sebagai berikut:

Remove Punctuation

Remove punctuation dapat disebut juga sebagai *clean text*. Data yang dihasilkan merupakan data-data yang telah dibersihkan dari karakter-karakter seperti tanda baca (,) (.) (/) dan angka (0,1,2,3,4,5,6,7,8,9). Adapun hasil dari *Remove punctuation* dapat dilihat pada Tabel 5 dan Tabel 6 berikut ini.

Tabel 2 Hasil data *Remove punctuation* Instagram

Dok 1	Gym nya kurang besar min
Dok2	lahan parkir nya luas gak min
Dok3	Selain fasilitas tempat ada fasilitas lain yg bisa mewadahi mahasiswa nya ga min
Dok4	unit kesehatan nya punya fasilitas lengkap min
Dok5	Minim nya fasilitas dalam ruangan kelas
Dok 7	Ada ketentuan buat peminjaman ruang podcast ga ya kemudian peminjaman ruangan sudah termasuk peminjaman alat alat ada dalam ruangan kah
Dok 8	perpustakaan nya bagus tapi saran saya lebih bagus lagi untuk filter ulang buku2 sudah lama ganti dengan buku2 keluaran baru
Dok 9	Uks nya belum ada perawat nya min
Dok 10	perpusnya bagus
Dok 11	Almamater
Dok 12	Mini gym buka hari apa aja min
Dok 13	ruang podcast bahas apa aja min
Dok 14	Mini market nya ada yg jaga ga ya ka Soalnya mau ke sana tapi tutup terus
Dok 15	Bahas fasilitas ada ProTeknik Mesin dong min

Tabel 3 Hasil Data *Remove punctuation* TikTok

Dok 1	Gym nya kurang besar min
Dok2	kelas masih pake bangku kayu tolong ganti min sama ac beberapa ruangan ada yg ga nginsemoga bisa membuat lebih nyaman lagi dlm belajar

```
# inialisasi stemming dengan module Sastrawi
factory = StemmerFactory()
stemmer = factory.create_stemmer()
```

Dok3	Mohon cek kamar mandi perempuan belakang masjid ya min suka ga ada airnya
Dok4	alat ruang gymnya bagus cocok untuk latihan pemula
Dok5	perpustakaan bagus banget min udah ganti AC aja lebih dingin

Case Folding

Case folding sebagai proses menyamakan bentuk huruf kapital menjadi huruf-huruf kecil. Hasilnya dapat dilihat pada Tabel 7 dan Tabel 8 berikut ini.

Tabel 4 Hasil Data *Case folding* Instagram

Dok1	gym nya kurang besar min
Dok2	lahan parkir nya luas gak min
Dok3	selain fasilitas tempat ada fasilitas lain yg bisa mewadahi mahasiswa nya ga min
Dok4	unit kesehatan nya punya fasilitas lengkap min
Dok5	minim nya fasilitas dalam ruangan kelas
Dok 7	ada ketentuan buat peminjaman ruang podcast ga ya kemudian peminjaman ruangan sudah termasuk peminjaman alat alat ada dalam ruangan kah
Dok 8	perpustakaan nya bagus tapi saran saya lebih bagus lagi untuk filter ulang buku2 sudah lama ganti dengan buku2 keluaran baru
Dok 9	uks nya belum ada perawat nya min
Dok 10	perpusnya bagus
Dok 11	Almamater
Dok 12	mini gym buka hari apa aja min
Dok 13	ruang podcast bahas apa aja min
Dok 14	mini market nya ada yg jaga ga ya ka soalnya mau ke sana tapi tutup terus
Dok 15	bahas fasilitas ada proteknik mesin dong min

Tabel 5 Hasil Data *Case folding* TikTok

Dok 1	gym nya kurang besar min
-------	--------------------------

	kelas masih pake bangku kayu tolong ganti min sama ac beberapa ruangan ada yg ga nginsemoga bisa membuat lebih nyaman lagi dlm belajar
Dok2	
	mohon cek kamar mandi perempuan belakang masjid ya min suka ga ada airnya
Dok3	
	alat ruang gymnya bagus cocok untuk latihan pemula
Dok4	
	perpustakaan bagus banget min udah ganti ac aja lebih dingin
Dok5	

Tokenizing

Tokenizing merupakan proses untuk memecah setiap kalimat menjadi sekumpulan kata tanpa tanda baca ataupun mengandung simbol-simbol tertentu. Pada kasus di atas, hasilnya sama seperti pada tabel 7 dan tabel 8.

TF IDF

Pada tahapan ini akan dilakukan proses perhitungan *term frequency* (TF), *document frequency* (DF), dan *inverse document frequency* (IDF). Adapun hasil pembobotan yang telah dilakukan yaitu sebagai berikut:

Tabel 9 Pembobotan TF-IDF

Term	TF				D F	W				
	K K	D 1	D 2	D 3		(Lo g D/d f)	KK	D 1	D 2	D 3
Gym	1	1	0	0	1	0.4 771 2	0.4 771 2	0	0	0
Fasilitas	0	0	1	0	1	0.4 771 2	0.4 771 2	1	0	0
Tempat	1	0	1	0	1	0.4 771 2	0.4 771 2	0	1	0
Ruang Podcast	1	0	0	1	1	0.4 771 2	0.4 771 2	0	1	1
UKS	1	0	1	1	1	0.4 771 2	0.4 771 2	0	0	1
Mini Market	1	0	0	0	1	0.4 771 2	0.4 771 2	0	0	0

c. Analisis Sentimen

Menghitung analisis sentimen dan hasil akhirnya (positif, negatif, atau netral)

melibatkan langkah yang harus diikuti. F1-score adalah perhitungan yang berguna karena menerima kedua *false positives* dan *false negatives* ke dalam pertimbangannya. Dalam konteks analisis sentimen, menghindari salah pengklasifikasian teks yang sebenarnya positif sebagai negatif (*false negative*) dan menghindari menyatakan teks yang sebenarnya negatif atau netral sebagai positif (*false positive*) keduanya penting. Tabel 10 berikut hasil dari perhitungan analisis sentimen pada *TikTok*:

Tabel 10 Hasil Perhitungan Analisis Sentimen

Label	F1-Score	Recall	Precision	Support
Positif	0,5	0,5	0,5	2
Negatif	0,6666666666	0,7	0,6	8
Netral	0	0	0	3

Pada tabel di atas, hasil *F1-score* pada sentimen positif sebesar 0.5 dan sentimen negative sebesar 0,6666666666666665

d. Klasifikasi Naïve Bayes

Naive bayes adalah algoritma klasifikasi yang sangat berguna dalam pemrosesan bahasa alami (NLP) dan analisis teks. Klasifikasi *Naïve Bayes* dilakukan dengan perhitungan probabilitas dan statistika. Data *Instagram* hasil perhitungan probabilitas dan statistika dapat dilihat dari data pada Tabel 11 berikut:

Tabel 11 Hasil Analisis Naïve Bayes Instagram

Kata	Positif	Negatif	Netral	P(X _i C _j)	PROBABILITAS		
					Positif	Negatif	Netral
gym	0	0	1	(0+1)/(4+1+14)	0.03 2258 1	0.07 142 9	0.058 82352 9
kurang	0	1	0		0.03 2258 1	0.14 285 7	0.029 41176 5
besar	1	0	0		0.06 4516 1	0.07 142 9	0.029 41176 5
min	1	0	0		0.06 4516 1	0.07 142 9	0.029 41176 5
lahan	0	0	1		0.03 2258 1	0.07 142 9	0.058 82352 9
parkir	0	0	1		0.03 2258 1	0.07 142 9	0.058 82352 9
lua	1	0	0		0.06	0.07	0.029

s				4516 1	142 9	41176 5
gak	0	1	0	0.03 2258 1	0.14 285 7	0.029 41176 5
min	0	0	1	0.03 2258 1	0.07 142 9	0.058 82352 9
sela in	0	0	1	0.03 2258 1	0.07 142 9	0.058 82352 9
fasi lita s	0	0	1	0.03 2258 1	0.07 142 9	0.058 82352 9
tem pat	0	0	1	0.03 2258 1	0.07 142 9	0.058 82352 9
ada	0	0	1	0.03 2258 1	0.07 142 9	0.058 82352 9
fasi lita s	1	0	0	0.06 4516 1	0.07 142 9	0.029 41176 5

Tabel 12 Hasil Naïve Bayes data *TikTok*

Kata	Positif	Negatif	Netral	P(Xi Cj)	PROBABILITAS		
					Positif	Negatif	Netral
gym	0	0	1	(0+1)/(4+1+14)	0.03225 8	0.07142 9	0.05882 4
kurang	0	1	0		0.03225 8	0.14285 7	0.02941 2
besar	1	0	0		0.06451 6	0.07142 9	0.02941 2
min	1	0	0		0.06451 6	0.07142 9	0.02941 2
kelas	0	0	1		0.03225 8	0.07142 9	0.05882 4
masih	0	0	1		0.03225 8	0.07142 9	0.05882 4
pakai	1	0	0		0.06451 6	0.07142 9	0.02941 2
bangk u	0	1	0		0.03225 8	0.14285 7	0.02941 2
kayu	0	1	0		0.03225 8	0.07142 9	0.05882 4
tolong	0	1	0		0.03225 8	0.07142 9	0.05882 4
ganti	0	0	1		0.03225 8	0.07142 9	0.05882 4
sama	0	0	1		0.03225 8	0.07142 9	0.05882 4
ac	0	0	1		0.03225 8	0.07142 9	0.05882 4
berapa	1	0	0		0.06451 6	0.07142 9	0.02941 2
ruanga n	0	0	1		0.03225 8	0.07142 9	0.05882 4
ada	1	0	0		0.06451 6	0.07142 9	0.02941 2
ga	0	1	0		0.06451 6	0.07142 9	0.05882 4
dingin	0	1	0		0.03225 8	0.07142 9	0.05882 4
semog a	0	0	1		0.03225 8	0.07142 9	0.05882 4
bisa	1	0	0		0.03225 8	0.14285 7	0.02941 2
buat	0	0	1		0.03225 8	0.07142 9	0.05882 4
lebih	1	0	0		0.03225 8	0.07142 9	0.05882 4

					8	9	4
--	--	--	--	--	---	---	---

e. Skenario Percobaan

Uji coba setiap data yang sudah di *crawling* atau sudah melewati tahap *crawling* data secara keseluruhan baik data training dan testing dengan menggunakan metode *cross validation*. Penggunaan data dari setiap url postingan *Instagram* dan *TikTok* pada akun FT-UMJ ada pada Gambar 4 berikut ini :

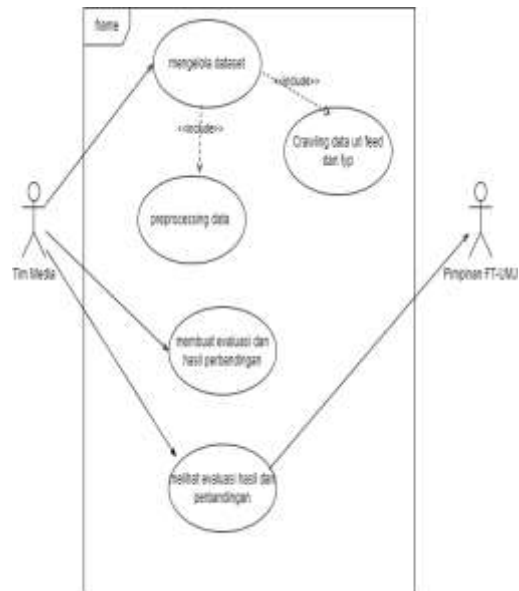


Percobaan 1	Testing	Training	Training	Training	Training
Percobaan 2	Training	Testing	Training	Training	Training
Percobaan 3	Training	Training	Testing	Training	Training
Percobaan 4	Training	Training	Training	Testing	Training
Percobaan 5	Training	Training	Training	Training	Testing

Gambar 4 Skenario Percobaan

Perancangan Desain Sistem

Perancangan desain sistem dilakukan dengan menentukan actor yang nantinya akan menggunakan. Dalam hal ini ada dua entitas, yaitu Tim Media dan Pimpinan FTUMJ, seperti pada gambar berikut ini :



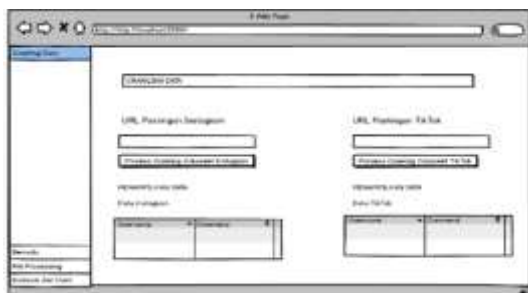
Gambar 5. use case diagram

Perancangan Tampilan Beranda Website

Desain atau *mockup* perancangan sistem dibuat agar gambaran yang dibuat dapat dipahami, dan dapat sebagai wadah komunikasi secara tepat dan cepat, dengan sistem yang telah dibuat. Beberapa *mockup* yang telah dibuat antara lain; 1)perancangan tampilan beranda, 2)perancangan tampilan *crawling* data, 3)perancangan tampilan *Preprocessing* data, 4) perancangan tampilan evaluasi data seperti pada Gambar 6, 7, 8 dan 9 berikut ini.



Gambar 6 Rancangan tampilan beranda

Gambar 7 Rancangan tampilan *crawling* dataGambar 8. Rancangan *preprocessing* data

Gambar 9. Rancangan evaluasi data

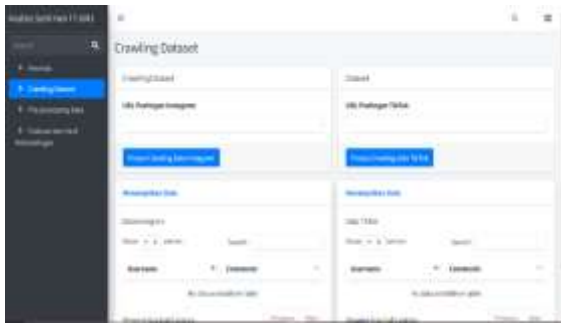
2. Hasil dan Pembahasan

Dari langkah analisis dan perancangan yang telah dilakukan di atas, selanjutnya dapat dilihat hasilnya. Pertama, dapat dilihat hasil tampilan beranda, *crawling* data, *Preprocessing* data, dan evaluasi data. Kedua, hasil analisis sentimen yang mengklasifikasikan teks menjadi tiga kategori utama: positif, negatif, dan netral. Ketiga, hasil visualisasi sentimen dan hasil klasifikasi pada akurasi.

Halaman beranda adalah halaman pertama, dan untuk memulainya dengan *klik* tombol mulai. Pada bagian pertama, hasilnya dapat dilihat seperti Gambar 10 berikut ini.

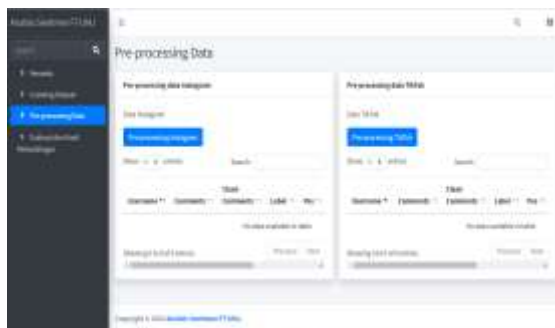


Gambar 10 Tampilan beranda
Halaman selanjutnya adalah *crawling* data set, seperti pada Gambar 11 berikut ini.



Gambar 11 Crawling dataset

Proses *crawling* data dilakukan dengan memasukkan link publikasi pada *feed* dan *fyp* *TikTok* dan *Instagram* di dalam *tools* lalu klik kolom proses *crawling* data *TikTok* dan proses *crawling* data *Instagram*. Proses *crawling* data dilakukan menggunakan *Apify* dan *API TikTok* dan *Instagram*. Setelah data berhasil diproses data akan terupdate otomatis ke dalam website, dan lanjut ke proses *Preprocessing* data seperti pada Gambar 12 berikut ini.



Gambar 12 Pre-processing data

Preprocessing data pada tampilan web dilakukan untuk membersihkan dan memformat data yang diterima dari pengguna melalui formulir atau interaksi dengan elemen tampilan web.

Setelah itu dilakukan evaluasi data seperti pada Gambar 13 berikut ini.



Gambar 13 Evaluasi data

Evaluasi data pada web merujuk pada proses memantau, menganalisis, dan

mengevaluasi data yang dikumpulkan dari pengguna atau sumber lain melalui aplikasi web.

Bagian kedua, hasil dari analisis sentimen dapat dilihat pada Gambar 14 berikut.



Gambar 14 Hasil Sentimen

Hasil analisis sentimen terhadap aplikasi *Instagram* dan *TikTok* dapat dilihat pada gambar di atas. Hasil Sentimen *Instagram*; positif 7, netral 6 dan negatif 13. Sementara hasil sentimen *TikTok*; positif 1, netral 1 dan negatif 3.

Kemudian pada bagian ketiga dilakukan perbandingan hasil visualisasi sentimen dengan melihat hasil akurasi pada aplikasi *Instagram* maupun *TikTok*.



Gambar 15 Hasil visualisasi sentiment

Visualisasi sentimen memungkinkan untuk dengan cepat melihat perbandingan antara sentimen positif, negatif, dan netral dari pengguna untuk masing-masing aplikasi. Pada Klasifikasi laporan *Instagram* diperoleh hasil akurasi sebesar 54% dan pada klasifikasi laporan *TikTok* diperoleh hasil akurasi sebesar 33%.



Gambar 16 Hasil klasifikasi pada akurasi

Hasil klasifikasi pada *Instagram* dan *TikTok* dilihat dari label (positif, negative dan netral), Nilai *F1-Score* yang menggabungkan *precision* (metrik yang mengukur kemampuan model untuk mendeteksi *instance* positif yang benar) dan *recall* (metrik yang mengukur kemampuan model untuk mendeteksi *instance* positif). Nilai *F-1 score*, *Recall* dan *Precision* berkisar antara 0 (terburuk) dan 1 (terbaik).

Pada gambar di atas dapat dilihat bahwa, *F-1 Score*, *Recall* dan *Precision* pada *Instagram* lebih baik untuk label negative dibandingkan pada *TikTok*. Sehingga dapat disimpulkan bahwa *Instagram* memiliki performa yang lebih baik dari *TikTok*.

3. Kesimpulan

Berdasarkan penelitian yang telah dilakukan, maka dapat diambil kesimpulan bahwa analisis sentimen tanggapan masyarakat dapat dilakukan menggunakan media sosial dengan aplikasi *Instagram* dan *TikTok* melalui unggahan *feed* dan *fyp* terhadap fasilitas kampus. Implementasi yang digunakan adalah *analisis sentimen lexicon base* dan *naïve bayes multinomial*. Hasil yang diperoleh adalah *F-1 Score*, *Recall* dan *Precision* pada *Instagram* lebih tinggi untuk label negative dibandingkan pada *TikTok*. Sehingga dapat disimpulkan bahwa *Instagram* dengan nilai akurasi 54% memiliki performa yang lebih baik dari *TikTok* yang memiliki nilai akurasi 33%. Sehingga, *Instagram* dapat direkomendasikan mendapatkan prioritas untuk dikembangkan lebih lanjut.

Saran

Penulis memberikan saran untuk menambahkan konten lain seperti komentar ataupun aplikasi lain misalnya menggunakan akun Twitter atau Youtube. Selain itu dapat

digunakan metode klasifikasi yang lainnya. Karena semakin bertambahnya perkembangan zaman maka semakin banyak pula model klasifikasi yang akan membantu proses penelitian

Ucapan Terima Kasih

Ucapan terima kasih Penulis berikan kepada Prodi Teknik Informatika yang telah memberikan dukungan dan inkind atas pelaksanaan penelitian ini.

Pembiayaan

Penelitian ini adalah penelitian mandiri yang tidak dibiayai oleh lembaga manapun.

Kontribusi Penulis

Penulis pertama berperan sebagai konseptor, merancang metodologi, merancang system dan mereviu serta mengedit hasil penelitian. Sementara penulis kedua berperan melakukan observasi dan wawancara, mengolah data, membuat aplikasi dan pengujian.

Daftar Pustaka

- Alfiah Zulqornain, J. a. (2021). "Analisis Sentimen Tanggapan Masyarakat Aplikasi TikTok Menggunakan Metode Naïve Bayes Dan Categorical Proportional Difference (CPD)." *j-ptiik*, 2886–90.
- Bulele, Y. N. (2020). "Analisis Fenomena Sosial Media Dan Kaum Milenial: Studi Kasus TikTok." *Conference on Business, Social Science and Innovation Technology* (pp. 565–72). Batam: Universitas Internasional Batam.
- Chaniago, R. H. (2020). "Budaya Komunikasi Dan Teknologi." *Journal of Chemical Information and Modeling*, 1689–99.
- Dwijayanti, A. (2022). "Efektivitas Penggunaan Media Sosial Sebagai Sarana Promosi Dan Pemasaran Pada UMKM Sablon Anggi Screen Di Era Digital." *Ikra-Ith Abdimas* 6, 68–75.

- Elfrida Cristella, Y. S. (2020). “*Analisis Sentimen Tentang Kasus Layanan Akademik Di Perguruan Tinggi*”. DKI Jakarta: FakultasFakultas Teknik Informatika. .
- Hermanto, H. A. (2020). “*Algoritma Klasifikasi Naive bayes Dan Support Vector Machine Dalam Layanan Komplain Mahasiswa*.” . Jakarta: JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer).
- Larasati, D. L. (2022). “Pengaruh Fasilitas Kampus Dan Kualitas Pelayanan Terhadap Kepuasan Mahasiswa Dengan Metode Sem The Effect of Campus Facilities and Service Quality on Student Satisfaction with SEM Method.”. *Jurnal Baut dan Manufaktur* , 2686–5351.
- Luthfiyah, A. (2023). *Perbandingan Sentimen Antara Aplikasi Tiktok Dan Instagram Terhadap Fasilitas Kampus Menggunakan Metode Klasifikasi Naive Bayes Multinomial*. Jakarta Pusat: Teknik Informatika UMJ.
- Rasdin, R. Y. (2021). “Fenomena Tik Tok Sebagai Media Komunikasi Edukasi.” . *Seminar Internasional Riksa Bahasa*, 227–35.
- Saifullah, S. M. (2017). "Analisa Terhadap Perbandingan Algoritma Decision Tree Dengan Algoritma Random Tree Untuk Pre-Processing Data.”. *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, 160.
- Sulistio, H. G. (2022). “Aspect-Based Sentiment Analysis Pada Ulasan ECommerce Dengan Metode Support Vector Machine Untuk Mendapatkan Informasi Sentimen Dari Beberapa Aspek. *Jurnal Infra*,
”<https://publication.petra.ac.id/index.php/teknik-infor>.
- Zhafira, D. F. (2021). “Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive bayes Dan Pembobotan TF-IDF Berdasarkan Komentar Pada Youtube.”. *Jurnal Sistem Informasi, Teknologi Informasi, dan Edukasi Sistem Informasi*, 55–63.