E-Jurnal Matematika Vol. 15(4), November 2026, pp. 302-312 ISSN: 2303-1751
DOI: https://doi.org/10.24843/MTK.2026.v15.i04.p534

ANALISIS SENTIMEN MENGGUNAKAN METODE NAIVE BAYES SEBAGAI STRATEGI PENGELOLAAN REPUTASI DIGITAL DAN CUSTOMER EXPERIENCE PADA CAFE GIGI SUSU

Diky Budiman[1§], Dewa Made Alit Adinugraha[2]

[1]Politeknik Internasional Bali [Email: budiman.diky@gmail.com]
[2]Politeknik Internasional Bali [Email: adinugraha.alit@gmail.com]
[§]*Corresponding Author*

ABSTRACT

Google Reviews serves as a vital source of customer insights to understand customer perception and support digital reputation management in the culinary business. This study aims to analyze customer sentiment across service, place, and price aspects at Café Gigi Susu using the Multinomial Naive Bayes method. The research employs a quantitative experimental approach using 760 cleaned reviews obtained through web scraping from Google Reviews. The data underwent aspect categorization with a priority hierarchy (Service > Place > Price), bilingual text preprocessing, stratified 80:20 data splitting, RandomOverSampler (ROS) class balancing on the training set, and TF-IDF feature extraction. The results show that 91.05% of the reviews were classified as positive and 8.95% as negative. The place aspect achieved the highest positive sentiment proportion (97.41%), whereas the service aspect accounted for the largest volume of negative reviews (33.82% of all complaints), and the price aspect exhibited the highest relative negative sentiment ratio (43.48%). Following class imbalance calibration with RandomOverSampler, the model achieved a test accuracy of 90.13%, increased minority negative-class recall from 0.00% to 50.00%, and attained an average 5-fold cross-validation accuracy of 96.12%. These findings demonstrate that calibrated aspect-based sentiment analysis effectively uncovers actionable customer insights to support service quality enhancement, customer experience optimization, and digital reputation management.

Keywords: *Sentiment Analysis, Google Reviews, Multinomial Naive Bayes, Digital Reputation, Customer Experience*

1. PENDAHULUAN

Perkembangan teknologi digital telah mengubah cara konsumen mencari informasi, mengevaluasi produk, dan mengambil keputusan pembelian. Salah satu perubahan yang paling signifikan adalah meningkatnya penggunaan platform ulasan daring (*online reviews*) sebagai sumber informasi utama sebelum konsumen melakukan transaksi. Dalam industri makanan dan minuman (*food and beverage industry*), khususnya bisnis kafe, ulasan pelanggan pada platform digital seperti *Google Review*s telah menjadi salah satu indikator penting yang mempengaruhi persepsi konsumen terhadap kualitas layanan, kualitas produk, dan reputasi suatu usaha. Pengelolaan *rating* dan *review* telah menjadi faktor strategis yang menentukan keberlangsungan dan daya saing bisnis (Mardikaningsih & Riski, 2026).

Rating dan *review* merupakan aset tidak berwujud yang dapat memengaruhi tingkat kepercayaan pelanggan, loyalitas konsumen, serta keputusan pembelian. Reputasi digital terbentuk melalui berbagai interaksi pelanggan di dunia maya, termasuk ulasan, komentar, rating, dan konten yang dihasilkan pengguna (*user-generated content*). Sebuah penelitian menunjukkan bahwa ulasan online memiliki pengaruh yang signifikan terhadap persepsi konsumen dan keputusan pembelian, terutama pada sektor *hospitality* dan layanan(Widiarta et al., 2024).

Bagi Cafe Gigi Susu, *Google Review*s

menjadi salah satu media yang paling banyak digunakan pelanggan untuk memberikan penilaian terhadap pengalaman mereka. Namun demikian, jumlah ulasan yang terus bertambah menyebabkan pemilik usaha mengalami kesulitan dalam menganalisis seluruh masukan pelanggan secara manual. Oleh karena itu, diperlukan pendekatan berbasis data melalui analisis sentimen (*sentiment analysis*). Analisis sentimen merupakan cabang dari *text mining* dan *Natural Language Processing* (NLP) yang bertujuan untuk mengidentifikasi, menganalisis, dan mengklasifikasikan opini, emosi, serta sikap yang diekspresikan melalui teks. Proses ini dilakukan dengan mengekstraksi informasi subjektif dari teks untuk menentukan kecenderungan sentimen yang terkandung di dalamnya, yang umumnya dikategorikan ke dalam sentimen positif, negatif, atau netral (Coletta et al., 2014; Julianto et al., 2022; Setiawan, 2025).

Berbagai penelitian terdahulu membuktikan efektivitas analisis sentimen. Penelitian yang dilakukan oleh Elistiana menerapkan algoritma Naive Bayes untuk menganalisis ulasan aplikasi Shopee (Elistiana et al., 2023). Pengaplikasian Multinomial Naive Bayes pada ulasan Pegipegi.com (Balit & Utomo, 2024). Sebuag penelitian membandingkan Naive Bayes, KNN, dan Logistic Regression pada ulasan hotel Novotel (Dhamma & Barus, 2025). penelitian lainnya menganalisis 2.891 ulasan wisatawan Pura Besakih (Sudiatmika et al., 2025). Penelitian pada ulasan Mie Gacoan Solo di Google Maps (Fatikhah et al., 2024). Penelitian lainnya mengkaji ulasan Kawah Ijen dengan menambahkan optimalisasi *oversampling* (Hizham et al., 2024). Azizah dkk. Melakukan pengujian pada ulasan Bukit Merese (Azizah et al., 2024). Sri & Saiful, menganalisis ulasan lokasi Google Maps (Sri & Saiful, 2024). Rizki & Prihartono, menganalisis ulasan rumah sakit dengan akurasi 84% (RIZKI & Prihartono, 2025). Serta Firdaus dkk, menerapkan *Aspect-Based Sentiment Analysis* (ABSA) pada ulasan restoran (Firdaus et al., 2024). Di samping itu, Inyang & White, menekankan pentingnya *Online Reputation Management* (ORM) terhadap kinerja usaha kecil (Inyang & White, 2026).

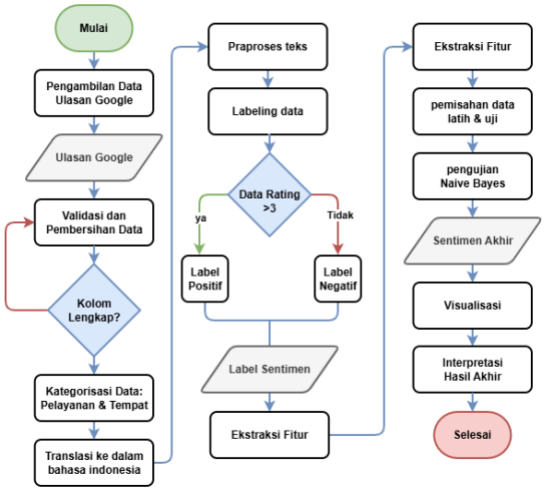
Meskipun demikian, kajian yang secara khusus mengintegrasikan klasifikasi sentimen berbasis aspek (servis, tempat, dan harga) dengan penanganan ketimpangan kelas (class imbalance) pada bisnis kafe lokal sebagai strategi pengelolaan reputasi digital dan customer experience masih terbatas. Penelitian ini menawarkan kebaruan konseptual dan metodologis berupa integrasi pemetaan sentimen multi-aspek, kalibrasi data tidak seimbang menggunakan RandomOverSampler, serta analisis kesenjangan antara skor rating dan isi ulasan pada Café Gigi Susu. Berdasarkan latar belakang tersebut, rumusan masalah dalam penelitian ini meliputi:

1. Bagaimana sentimen pelanggan terhadap aspek servis, tempat, dan harga pada Café Gigi Susu berdasarkan Google Reviews yang dianalisis menggunakan metode Multinomial Naive Bayes?
2. Bagaimana performa dan stabilitas model Multinomial Naive Bayes dalam mengklasifikasikan sentimen ulasan pelanggan setelah penerapan penanganan class imbalance menggunakan RandomOverSampler dan pengujian 5-Fold Cross-Validation?

Tujuan penelitian ini adalah menganalisis polaritas sentimen pelanggan terhadap aspek servis, tempat, dan harga pada Café Gigi Susu menggunakan algoritma *Multinomial Naive Bayes* yang dikalibrasi dengan metode *RandomOverSampler*, guna menghasilkan wawasan pelanggan (*customer insights*) yang mendukung pengelolaan reputasi digital serta peningkatan *customer experience*.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan penerapan algoritma *Multinomial Naive Bayes Classifier* (NBC), merupakan metode klasifikasi statistik berbasis Teorema Bayes yang memprediksi probabilitas suatu data masuk ke kelas tertentu dengan mengasumsikan bahwa setiap atribut saling independen (Pebdika et al., 2023). Subjek penelitian adalah ulasan pelanggan Café Gigi Susu yang berlokasi di Canggu, Kabupaten Badung, Bali, yang dikumpulkan dari platform *Google Reviews*. Untuk menjamin validitas pemodelan dan mengatasi ketimpangan distribusi kelas (*class imbalance*), rancangan penelitian mengintegrasikan ekstraksi fitur *Term Frequency–Inverse Document Frequency* (TF-IDF), penyeimbangan data latih menggunakan *RandomOverSampler* (ROS), pembagian dataset secara konsisten dengan rasio 80:20, serta pengujian stabilitas generalisasi model menggunakan teknik *5-Fold Cross-Validation*.



Gambar 2.1 Alur Program Analisis Sentimen

2.1 Alur Penelitian

Tahap pertama dimulai dari proses pengambilan data ulasan Google. Data yang digunakan berasal dari *Google Review* dengan atribut utama berupa *reviewed, name, stars, dan text*. Kolom stars digunakan sebagai indikator rating pengguna, sedangkan kolom *text* menjadi sumber utama dalam proses analisis sentimen berbasis teks. Setelah data diperoleh, tahap berikutnya adalah validasi dan pembersihan data. Tahap ini bertujuan untuk memastikan bahwa data yang digunakan telah memenuhi kebutuhan analisis. Program melakukan pemeriksaan terhadap kelengkapan kolom wajib, yaitu *reviewed, name, stars, dan text.* Apabila terdapat kolom yang tidak lengkap, maka proses diarahkan kembali ke tahap validasi agar data diperbaiki terlebih dahulu. Namun, apabila seluruh kolom telah tersedia, proses dilanjutkan ke tahap kategorisasi. Selain validasi kolom, pembersihan data juga mencakup penghapusan data kosong, penghapusan ulasan tanpa teks, konversi rating menjadi numerik, serta penghapusan data duplikat. Tahap ini penting karena analisis sentimen berbasis teks memerlukan data komentar yang memiliki kandungan bahasa, sehingga data tanpa ulasan tidak dapat memberikan informasi yang cukup untuk proses klasifikasi.

Tahap selanjutnya adalah kategorisasi data berdasarkan aspek pelayanan dan tempat. Pada tahap ini, ulasan dikelompokkan ke dalam tiga kategori, yaitu Servis, Tempat, dan Tidak Terkategori. Kategorisasi dilakukan menggunakan pendekatan berbasis kata kunci. Kata-kata seperti "pelayanan", "layanan", "staff", "ramah", "kasir", dan "pesanan" digunakan untuk mengidentifikasi ulasan yang berkaitan dengan aspek servis. Sementara itu, kata-kata seperti "tempat", "suasana", "ambience", "nyaman", "lokasi", "interior", dan "parkir" digunakan untuk mengidentifikasi ulasan yang berkaitan dengan aspek tempat. Setelah kategorisasi, data ulasan diterjemahkan ke dalam bahasa Indonesia. Tahap translasi ini diperlukan karena *Google Review* dapat ditulis dalam berbagai bahasa, seperti bahasa Indonesia dan bahasa Inggris. Dengan menerjemahkan seluruh ulasan ke dalam bahasa Indonesia, proses praproses teks dan analisis kata menjadi lebih seragam. Setelah proses translasi, data memasuki tahap praproses teks. Pada tahap ini, program mengubah seluruh teks menjadi huruf kecil, menghapus tautan, simbol, tanda baca, angka, karakter khusus, serta menghapus stopword menggunakan pustaka Sastrawi. Tujuan dari tahap ini adalah menghasilkan teks bersih yang lebih siap digunakan dalam proses ekstraksi fitur dan pemodelan sentimen.

Tahap berikutnya adalah pelabelan data berdasarkan rating. Keputusan ini menjadi dasar dalam pembentukan label sentimen awal. Apabila rating ulasan lebih dari 3, maka data diberi label Positif. Sebaliknya, apabila rating bernilai 3 atau kurang, maka data diberi label Negatif. Pendekatan ini disebut sebagai pelabelan otomatis berbasis rating atau weak labeling, karena rating pengguna diasumsikan mencerminkan tingkat kepuasan terhadap destinasi yang diulas. Setelah label sentimen terbentuk, proses dilanjutkan ke tahap ekstraksi fitur. Pada tahap ini, teks yang telah dibersihkan diubah menjadi representasi numerik agar dapat diproses oleh algoritma pembelajaran mesin.

Selanjutnya dilakukan pemisahan data latih dan data uji. Dalam program, dataset dibagi dengan proporsi 80:20, yaitu 80% data digunakan sebagai data latih dan 20% data digunakan sebagai data uji. Pembagian ini bertujuan untuk melatih model pada sebagian besar data dan menguji performa model pada data yang belum pernah dilihat sebelumnya. Setelah pemisahan data dilakukan, model klasifikasi dibangun menggunakan algoritma *Multinomial Naive Bayes*. Algoritma ini dipilih karena sesuai untuk klasifikasi teks. Setelah proses klasifikasi, flowchart menunjukkan tahap penentuan sentimen akhir. Pada program, hasil prediksi Naive Bayes dibandingkan dengan label berbasis rating. Apabila label rating dan hasil prediksi model sama, maka data diberi

status "Sesuai". Namun, apabila terdapat perbedaan antara label rating dan prediksi model, maka data diberi status "Konflik". Keberadaan status konflik ini penting karena tidak semua rating selalu mencerminkan isi komentar secara sempurna. Sebagai contoh, pengguna dapat memberikan rating tinggi tetapi tetap menyampaikan kritik dalam teks ulasan, atau sebaliknya. Oleh karena itu, sentimen akhir tetap mempertimbangkan label berbasis rating, sementara prediksi model digunakan sebagai pembanding untuk melihat konsistensi antara rating dan isi teks.

Tahap berikutnya dilakukan visualisasi hasil untuk memudahkan interpretasi data secara deskriptif. Program menghasilkan beberapa bentuk visualisasi, seperti diagram proporsi sentimen keseluruhan, grafik sentimen berdasarkan kategori review, confusion matrix, *Word Cloud* kategori Servis dan Tempat, serta visualisasi sepuluh kata dominan pada sentimen positif dan negatif. Visualisasi ini membantu memperlihatkan pola utama dalam ulasan pelanggan. Misalnya, kata-kata seperti "enak", "bagus", "ramah", "nyaman", "makanan", "*croissant*", dan "kopi" banyak muncul pada sentimen positif. Hal ini menunjukkan bahwa kepuasan pelanggan terhadap Gigi Susu dibentuk oleh kualitas produk, kenyamanan tempat, dan keramahan pelayanan.

Pembentukan model klasifikasi untuk memperoleh nilai akurasi diawali dengan penentuan parameter yang tepat agar model mampu menghasilkan performa yang optimal. Pemilihan parameter yang kurang sesuai dapat memengaruhi kualitas sistem, baik dalam meningkatkan maupun menurunkan tingkat akurasi klasifikasi. Oleh karena itu, pada penelitian ini dilakukan pembagian *dataset* untuk memperoleh parameter terbaik. Pembagian *dataset* dilakukan setelah seluruh tahapan *preprocessing* selesai, sehingga data terbagi menjadi dua kelompok, yaitu *training data* dan *testing data*, dengan komposisi sebesar lima banding tiga. Kedua kelompok data tersebut digunakan dalam proses klasifikasi menggunakan metode *Naive Bayes Classifier (NBC)*. *Training data* digunakan untuk membangun model melalui perhitungan probabilitas *prior* sebagai dasar pembelajaran, sedangkan *testing data* dimanfaatkan untuk mengevaluasi model dengan menghitung probabilitas setiap kelas sentimen. Hasil klasifikasi ditentukan berdasarkan kelas yang memiliki nilai probabilitas tertinggi, sehingga setiap dokumen akan diberi label sentimen positif, netral, atau negatif. Untuk memperoleh parameter yang paling optimal, dilakukan serangkaian pengujian menggunakan teknik *K-Fold Cross Validation* dengan membandingkan hasil klasifikasi pada setiap pembagian data yang berbeda (Kamal & Ratnasari, 2021).

2.2 Formulasi Matematis Multinomial Naive Bayes

Representasi numerik setiap kata (*term*) *t* di dalam dokumen ulasan *d* dihitung menggunakan metode pembobotan *Term Frequency–Inverse Document Frequency* (TF-IDF) melalui Persamaan (1) dan Persamaan (2):

$$w_{t,d} = \mathrm{tf}(t,d) \times \mathrm{idf}(t) \quad (1)$$

$$\mathrm{idf}(t) = \log\left(\frac{N}{\mathrm{df}(t)}\right) + 1 \quad (2)$$

dengan tf(*t*, *d*) menunjukkan frekuensi kemunculan term *t* pada dokumen *d*, *N* adalah jumlah total dokumen dalam korpus pelatihan, dan df(*t*) merupakan jumlah dokumen yang memuat term *t*.

Selanjutnya, algoritma *Multinomial Naive Bayes* (MNB) menghitung probabilitas posterior suatu dokumen ulasan $d = (t_1, t_2, \ldots, t_n)$ termasuk ke dalam kelas sentimen biner $c \in$ {Positif, Negatif} berdasarkan Teorema Bayes dengan asumsi independensi bersyarat antar-fitur (*conditional independence*), sebagaimana dirumuskan pada Persamaan (3):

$$\mathrm{P}(c \mid d) = \frac{\mathrm{P}(c)\ \prod_{i=1}^{n} \mathrm{P}(t_i \mid c)^{w_{i,d}}}{\mathrm{P}(d)} \quad (3)$$

Karena nilai probabilitas bukti P(*d*) bernilai konstan untuk seluruh kelas pada dokumen yang sama, penentuan kelas prediksi optimal cukup memaksimalkan pembilang pada Persamaan (3). Estimasi probabilitas prior kelas diperoleh dari rasio jumlah dokumen kelas *c* terhadap total dokumen latih pasca-penyeimbangan *RandomOverSampler* (*N*train), sesuai Persamaan (4):

$$\widehat{\mathrm{P}}(c) = \frac{N_c}{N_{\mathrm{train}}} \quad (4)$$

Sementara itu, probabilitas kondisional (likelihood) fitur diestimasi menggunakan *Laplace Smoothing* (*additive smoothing*) dengan parameter $\alpha = 1$ untuk mencegah munculnya probabilitas nol (*zero-frequency problem*) ketika terdapat kata pada data uji yang tidak ditemukan pada kelas *c* saat pelatihan, sebagaimana ditunjukkan pada Persamaan (5):

$$\hat{\mathrm{P}}(t_i \mid c) = \frac{W_{t_i,c} + \alpha}{\sum_{t\prime \in V} W_{t\prime,c} + \alpha|V|} \quad (5)$$

dengan W menyatakan total bobot TF-IDF term pada seluruh dokumen pelatihan berkelas c, V merupakan himpunan kosakata unik (*vocabulary*), dan $|V|$ adalah ukuran dimensi kosakata. Guna menghindari kesalahan komputasi *numerical underflow* akibat perkalian nilai probabilitas yang sangat kecil pada dokumen ulasan yang panjang, aturan keputusan akhir *Maximum A Posteriori* (MAP) ditransformasikan ke dalam ruang logaritma sesuai Persamaan (6):

$$\hat{c} = \arg\max \left[\log \hat{\mathrm{P}}(c) + \sum_{i=1}^{n} w_{i,d} \log \hat{\mathrm{P}}(t_i \mid c) \right] \quad (6)$$

Kelas biner $c \in$ {Positif, Negatif} yang menghasilkan nilai log-posterior tertinggi ditetapkan sebagai kelas sentimen prediksi akhir pada ulasan tersebut.

2.3 Perhitungan Performa

Tahap akhir dalam penelitian ini adalah melakukan evaluasi terhadap kinerja model klasifikasi dengan menghitung nilai *accuracy*, *precision*, *recall*, dan *F1-score* menggunakan metode *Confusion Matrix*. *Confusion Matrix* merupakan salah satu metode evaluasi yang digunakan untuk mengukur tingkat ketepatan model dalam mengklasifikasikan data ke dalam setiap kelas sentimen (Arminda et al., 2023). Melalui metode ini, performa model dapat dianalisis berdasarkan jumlah prediksi yang benar maupun yang salah pada masing-masing kategori. Konsep dasar penggunaan *Confusion Matrix* dalam penelitian ini disajikan pada Tabel 2.1

Tabel 2.1 *Confusion Matrix*

Confusion Matrix		*True Class*	
		Positive	*Negative*
Predicted Class	*Positive*	*True Positive (TP)*	*False Positive (FP)*
	Negative	*False Negative (FN*	*True Negative (TN)*

Berdasarkan komponen pada Tabel 2.1, dihitung empat metrik evaluasi utama, yakni *Accuracy*, *Precision*, *Recall*, dan *F1-Score* menggunakan Persamaan (7) hingga Persamaan (10):

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (7)$$

$$\text{Precision}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FP}_c} \quad (8)$$

$$\text{Recall}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FN}_c} \quad (9)$$

$$\text{F1} - \text{Score}_c = 2 \times \frac{\text{Precision}_c \times \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c} \quad (10)$$

Accuracy mengukur proporsi ketepatan klasifikasi secara keseluruhan. *Precision* mengukur tingkat relevansi prediksi pada suatu kelas, *Recall* mengukur sensitivitas model dalam mengenali seluruh sampel aktual suatu kelas (terutama kelas Negatif yang berjumlah minoritas), sedangkan *F1-Score* memadukan nilai *Precision* dan *Recall* melalui rata-rata harmonis guna memberikan ukuran evaluasi yang berimbang pada distribusi data tidak seimbang.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengolahan Dataset Awal

Penelitian ini menggunakan *dataset Google Review* pada destinasi Gigi Susu. *Dataset* terdiri atas beberapa atribut utama, yaitu `*reviewerId*`, `*name*`, `*stars*`, dan `*text*`. Kolom `*stars*` digunakan sebagai indikator rating pengguna, sedangkan kolom `*text*` digunakan sebagai sumber utama analisis sentimen berbasis teks. Sebelum dilakukan analisis sentimen, data terlebih dahulu melewati tahap pembersihan awal. Tahap ini dilakukan dengan menghapus ulasan yang tidak memiliki komentar, menghapus data kosong, mengubah nilai rating menjadi numerik, serta menghilangkan data duplikat.

Dataset menunjukkan bahwa sekitar 76% data dapat digunakan dalam analisis, sedangkan sekitar 24% data dieliminasi. Penghapusan data tanpa komentar merupakan langkah penting karena analisis sentimen berbasis teks membutuhkan kandungan linguistik yang dapat diproses. Apabila data tanpa komentar tetap dipertahankan, maka proses klasifikasi sentimen dapat menjadi bias karena model tidak memiliki informasi tekstual yang cukup untuk menentukan kecenderungan sentimen pengguna.

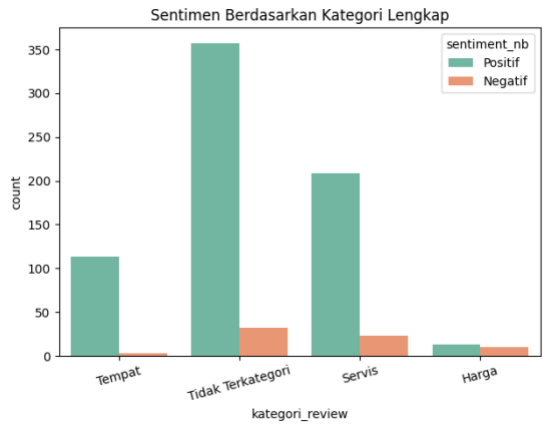
3.2 Kategorisasi Aspek Ulasan dan Aturan Prioritas Kata Kunci

Untuk menggali dimensi persepsi pelanggan secara lebih terperinci, seluruh 760 ulasan bersih dikelompokkan berbasis kata

kunci (*keyword-based categorization*) ke dalam tiga aspek operasional utama, yakni **Servis**, **Tempat**, dan **Harga** serta satu kategori **Tidak Terkategori** bagi ulasan umum yang tidak memuat kata kunci eksplisit dari ketiga aspek tersebut. Penambahan aspek Harga melengkapi dimensi layanan dan atmosfer fisik agar evaluasi pengalaman pelanggan mencerminkan persepsi nilai ekonomis (*value-for-money*) secara utuh. Mengingat ulasan pelanggan kerap bersifat multidimensi dan menyebut lebih dari satu aspek dalam satu komentar, diterapkan aturan hierarki prioritas (*tie-breaker*) dengan urutan Servis > Tempat > Harga. Urutan prioritas ini ditetapkan berdasarkan tingkat sensitivitas interaksi operasional dan urgensi pemulihan layanan (*service recovery*), di mana friksi interaksi staf memerlukan respons manajerial paling segera dibandingkan penilaian elemen fisik maupun persepsi kewajaran harga. Hasil pemetaan distribusi aspek dan perbandingan sentimen berdasarkan klasifikasi model *Multinomial Naive Bayes* (*sentiment_nb*) ditunjukkan pada Tabel 3.2 dan divisualisasikan pada Gambar 3.1.

Tabel 3.2 Distribusi Ulasan dan Komparasi Sentimen Antar-Aspek Berdasarkan Model *Multinomial Naive Bayes*

Kategori Aspek	Proporsi Dataset
Servis	30,53%
Tempat	15,26%
Harga	3,03%
Tidak Terkategori	51,18%
Total Keseluruhan	100,00%



Gambar 3.1 Proporsi Sentimen Berdasarkan Kategori Aspek dan Distribusi Sentimen Keseluruhan

Secara kuantitatif Tabel 3.2 memperlihatkan dua temuan struktural yang penting. Pertama, kelompok Tidak Terkategori mencakup 389 ulasan (51,18%). Pemeriksaan pada korpus menunjukkan bahwa kelompok ini sebagian besar berisi ulasan singkat atau penilaian yang berfokus langsung pada atribut produk kuliner (*pastry*, *almond croissant*, kopi, *matcha*) tanpa menyinggung kata kunci pelayanan, tempat, maupun harga secara eksplisit (misalnya *"Kopi enak"* atau *"Everything is delicious, food, coffee, and pastries 10/10"*). Hal ini mengindikasikan bahwa pendekatan pencocokan kata kunci leksikal efektif untuk menjaring opini berorientasi aspek spesifik, tetapi memiliki keterbatasan dalam menangkap variasi ekspresi implisit sehingga membuka ruang pengembangan *Aspect-Based Sentiment Analysis* (ABSA) atau *topic modeling* pada riset lanjutan. Kedua, di antara tiga aspek terdefinisi, Servis menempati porsi terbesar dengan 232 ulasan (30,53%), diikuti oleh aspek Tempat sebanyak 116 ulasan (15,26%), dan aspek Harga sebanyak 23 ulasan (3,03%). Tingginya frekuensi kemunculan aspek Servis menegaskan bahwa kualitas interaksi antara pelanggan dan staf merupakan titik sentuh (*touchpoint*) yang paling intens dievaluasi pengunjung kafe.

3.3 Analisis Sentimen Keseluruhan dan Komparasi Antar-Aspek

Berdasarkan hasil klasifikasi akhir model *Multinomial Naive Bayes* terkalibrasi (*sentiment_nb*), persepsi pelanggan terhadap Café Gigi Susu secara keseluruhan didominasi oleh sentimen positif sebanyak 692 ulasan (91,05%), sedangkan sentimen negatif tercatat sebanyak 68 ulasan (8,95%) (Gambar 3.1). Apabila dibandingkan secara komparatif antar-aspek pada Tabel 3.2, terlihat kontras yang tajam antara kekuatan utama destinasi dan titik kerentanan operasionalnya:

1. Keunggulan Kompetitif pada Aspek Tempat: Aspek Tempat mencatat tingkat kepuasan tertinggi dengan 97,41% sentimen positif (113 ulasan) dan hanya 2,59% sentimen negatif (3 ulasan), atau sekadar menyumbang 4,41% dari total keluhan. Temuan ini mengonfirmasi bahwa desain interior, atmosfer ruang (*ambience*), serta daya tarik visual yang fotogenik merupakan keunggulan kompetitif utama Café Gigi Susu dalam membentuk kesan positif pengunjung.
2. Volume Keluhan Tertinggi pada Aspek Servis: Meskipun aspek Servis memperoleh sentimen positif sebesar 90,09% (209 ulasan), aspek ini menghasilkan jumlah

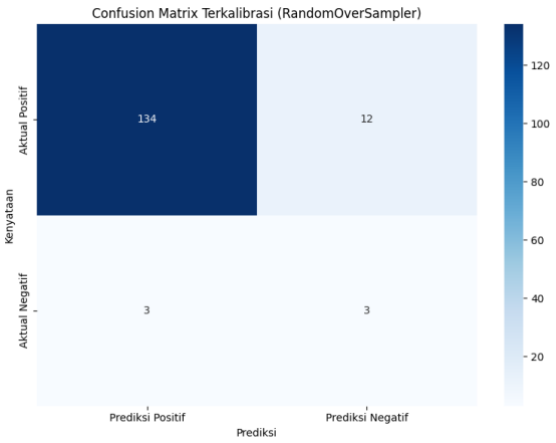
ulasan negatif absolut tertinggi di antara kategori terdefinisi, yakni 23 ulasan (9,91%) atau 33,82% dari total sentimen negatif. Analisis isi ulasan menunjukkan bahwa kritik pada aspek servis umumnya dipicu oleh keramaian pada jam sibuk (*peak hours*), waktu tunggu antrean yang panjang, keterbatasan fleksibilitas staf terhadap permintaan modifikasi pesanan, serta koordinasi petugas parkir dan keamanan (*security*).

3. Sensitivitas Keluhan Relatif Tertinggi pada Aspek Harga: Aspek Harga memperlihatkan pola polarisasi sentimen yang paling mencolok. Dari 23 ulasan yang membahas harga secara eksplisit, 10 ulasan (43,48%) terklasifikasi bersentimen negatif dan 13 ulasan (56,52%) bersentimen positif, sehingga menyumbang 14,71% dari seluruh ulasan negatif meskipun porsi ulasannya hanya 3,03% dari dataset. Tingginya rasio sentimen negatif pada aspek Harga mengindikasikan bahwa sebagian pelanggan menilai harga menu relatif mahal (*pricey* atau *overprice*) apabila dibandingkan dengan ukuran porsi, waktu antre, atau ekspektasi rasa produk yang dibentuk oleh popularitas media sosial.

Lebih jauh, analisis komparatif antara label berbasis rating (*sentiment_star*) dan hasil prediksi tekstual model *Multinomial Naive Bayes* (*sentiment_nb*) mengungkap tingkat kesesuaian (*agreement*) sebesar 93,95% (714 ulasan) dan tingkat konflik (*discrepancy*) sebesar 6,05% (46 ulasan). Sebanyak 43 dari 46 kasus konflik merupakan ulasan dengan peringkat bintang tinggi (Bintang 4 dan 5) yang terdeteksi memiliki muatan sentimen Negatif oleh model teks, tersebar pada kategori *Tidak Terkategori* (23 ulasan), *Servis* (10 ulasan), *Harga* (8 ulasan), dan *Tempat* (2 ulasan). Pola ini merefleksikan fenomena *hedged praise* atau bias kesopanan pelanggan: pengunjung tetap memberikan peringkat 4–5 bintang karena puas dengan estetika tempat atau rasa *pastry*, tetapi secara eksplisit menuliskan kritik terkait harga yang kurang terjangkau (*"its not affordable, but worth to buy"*), lamanya antrean foto dan pemesanan, atau kekhawatiran terhadap *overhype*. Dengan demikian, klasifikasi berbasis teks mampu menangkap sinyal peringatan dini (*early warning signals*) yang kerap tersembunyi di balik skor rating bintang yang tinggi.

3.4 Evaluasi Performa Model dan Penanganan Class Imbalance

Pengujian klasifikasi dilakukan menggunakan algoritma *Multinomial Naive Bayes* dengan pembagian dataset secara konsisten sebesar 80:20, yaitu 608 ulasan sebagai data latih dan 152 ulasan sebagai data uji (terdiri atas 146 data uji aktual Positif dan 6 data uji aktual Negatif). Untuk mengatasi bias prediksi ke kelas mayoritas akibat ketimpangan distribusi kelas (96,32%: 3,68%), penelitian ini menerapkan teknik penyeimbangan data *RandomOverSampler* (ROS) secara langsung pada partisi data latih sebelum proses pelatihan model dijalankan. Hasil evaluasi pada data uji disajikan melalui *Confusion Matrix* terkalibrasi pada Gambar 3.2, sedangkan perbandingan metrik kinerja sebelum dan sesudah penerapan ROS dirangkum pada Tabel 3.3.

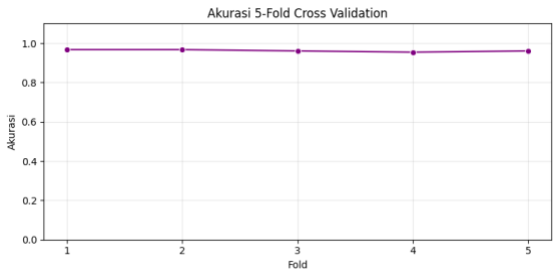


Gambar 3.2 *Confusion Matrix* Model *Multinomial Naive Bayes* Terkalibrasi (*RandomOverSampler*)

Tabel 3.3 Perbandingan Evaluasi Performa Model Sebelum dan Sesudah Penerapan *RandomOverSampler* (ROS)

Skenario Model	Metrik Evaluasi	Kelas Positif (n=146)	Kelas Negatif (n=6)	Macro Average
MNB (Tanpa ROS)	*Precision*	95,39%	0,00%	47,70%
Accuracy: 95,39%	*Recall*	100,00%	0,00%	50,00%
	F1-Score	97,64%	0,00%	48,82%
MNB Terkalibrasi (+ ROS)	*Precision*	97,81%	20,00%	58,91%
Accuracy: 90,13%	*Recall*	91,78%	**50,00%**	**70,89%**
	F1-Score	94,70%	**28,57%**	**61,64%**

Berdasarkan Gambar 3.2, dari total 152 data uji, model terkalibrasi menghasilkan 134 True Positive (TP), 12 False Negative (FN), 3 False Positive (FP), dan 3 True Negative (TN). Analisis komparatif pada Tabel 3.3 menunjukkan dampak substantif penerapan ROS terhadap perilaku klasifikasi model. Pertama, pada skenario *baseline* tanpa penyeimbangan data, model memang mencatatkan akurasi semu sebesar 95,39%, tetapi gagal total mengenali ulasan negatif (*Recall* = 0,00% dan *F1-Score* = 0,00%) karena seluruh data uji diprediksi sebagai Positif. Setelah penerapan *RandomOverSampler*, nilai Recall kelas Negatif berhasil meningkat drastis dari 0,00% menjadi 50,00% dengan *F1-Score* meningkat menjadi 28,57%, sekaligus mendongkrak *Macro Average Recall* dari 50,00% menjadi 70,89%. Kedua, peningkatan sensitivitas deteksi keluhan diiringi dengan penurunan moderat akurasi keseluruhan sebesar 5,26 poin persentase (dari 95,39% menjadi 90,13%) serta nilai *Precision* kelas Negatif sebesar 20,00% akibat munculnya 12 prediksi *False Negative* (ulasan positif yang mengandung kata bernuansa kritis seperti *"overrated"*, *"pricey"*, atau *"antri"* sehingga terdeteksi sebagai negatif oleh model). Dalam konteks pengelolaan reputasi digital bisnis perhotelan dan kuliner, *trade-off* ini sangat menguntungkan secara praktis karena biaya manajerial akibat melewatkan keluhan pelanggan nyata jauh lebih besar daripada meninjau ulang ulasan positif yang bernada kritis.



Gambar 3.3 Grafik Stabilitas Akurasi Pengujian *5-Fold Cross-Validation*

Untuk memastikan bahwa performa model tidak bergantung pada satu partisi pembagian data tertentu, dilakukan validasi silang menggunakan *5-Fold Cross-Validation* (Gambar 3.3). Hasil pengujian menunjukkan rata-rata akurasi sebesar 96,12% dengan kurva performa lintas *fold* 1 hingga 5 yang bergerak sangat stabil pada rentang 95,5%–96,8% tanpa fluktuasi ekstrem. Konsistensi lintasan grafik pada Gambar 3.3 membuktikan bahwa kombinasi representasi fitur TF-IDF, penyeimbangan *RandomOverSampler*, dan klasifikasi *Multinomial Naive Bayes* memiliki kapasitas generalisasi yang andal untuk memetakan polaritas opini ulasan pelanggan.

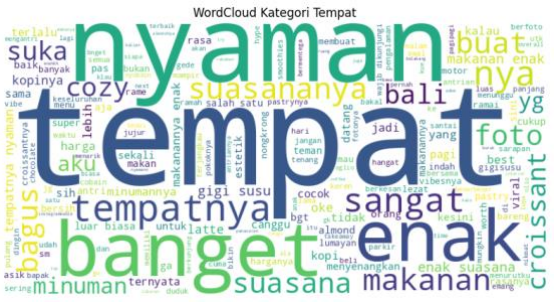
3.5 Analisis Kata Dominan (WordCloud) Berdasarkan Aspek

Untuk memperdalam interpretasi semantik di balik angka klasifikasi, dilakukan visualisasi frekuensi kata dominan (*WordCloud*) pada masing-masing kategori aspek, yakni Aspek Service (Gambar 3.4), Aspek Tempat (Gambar 3.5), dan Aspek Harga (Gambar 3.6).



Gambar 3.4 *WordCloud* Kategori *Service* (Servis)

Pada Gambar 3.4, korpus ulasan aspek *Service* didominasi oleh leksikon apresiatif seperti *"ramah"*, *"enak"*, *"baik"*, *"bagus"*, *"staff"*, *"pelayanan"*, dan *"cepat"*. Ukuran kata *"ramah"* dan *"staff"* yang paling menonjol mengindikasikan bahwa keramahtamahan serta sikap personal pramusaji merupakan pendorong utama kepuasan layanan di Café Gigi Susu. Namun, kemunculan kata penghubung evaluatif seperti *"tidak"*, *"terlalu"*, serta kata terkait alur pemesanan seperti *"pesanan"* dan *"antri"* mencerminkan titik friksi operasional yang menyumbang 23 ulasan negatif pada aspek ini.



Gambar 3.5 *WordCloud* Kategori Tempat

Pada Gambar 3.5, *WordCloud* aspek Tempat memperlihatkan dominasi kuat pada kata *"nyaman"*, *"tempat"*, *"suasananya"*, *"bagus"*, *"santai"*, *"cozy"*, dan *"foto"*. Dominasi kata *"nyaman"* dan *"cozy"* berpadu dengan kata *"foto"* dan *"estetik"* menegaskan bahwa Café Gigi Susu tidak sekadar dipersepsikan sebagai tempat konsumsi makanan dan minuman, melainkan sebagai ruang pengalaman (*experiential space*) yang mendukung aktivitas bersantai sekaligus dokumentasi visual bagi wisatawan. Keselarasan atribut fisik dan atmosfer inilah yang menjaga tingkat sentimen positif aspek Tempat tetap bertengger pada angka 97,41%.



Gambar 3.6 *WordCloud* Khusus Aspek Harga

Sementara itu, Gambar 3.6 (*WordCloud* khusus aspek Harga) memperlihatkan dinamika persepsi nilai yang bersifat dua kutub (*polarized*). Di satu sisi, kata kunci evaluasi nilai positif seperti *"worth"*, *"good"*, *"enak"*, *"croissant"*, dan *"coffee"* tampil dominan, yang menandakan bahwa pelanggan bersedia membayar lebih apabila kualitas produk utama (*signature pastry* dan kopi) memenuhi ekspektasi rasa. Di sisi lain, kemunculan kata-kata kritis seperti *"pricey"*, *"overprice"*, *"agak"*, *"terlalu"*, *"hype"*, dan *"fomo"* memperlihatkan bahwa sebagian pengunjung mempertanyakan kesepadanan harga dengan pengalaman keseluruhan (*value-for-money*), terutama ketika ekspektasi tinggi akibat viralitas media sosial berbenturan dengan antrean panjang atau porsi hidangan.

4. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Hasil penelitian ini memperkuat sekaligus memperluas temuan sejumlah studi analisis sentimen pada sektor kuliner dan perhotelan. Keandalan algoritma *Multinomial Naive Bayes* dalam mengklasifikasikan ulasan digital berbahasa campuran (Indonesia–Inggris) dengan rata-rata akurasi *cross-validation* sebesar 96,12% sejalan dengan temuan Balit dan Utomo (2024) pada ulasan platform perjalanan serta Fatikhah et al. (2024) dan Sudiatmika et al. (2025) pada ulasan destinasi kuliner dan wisata di Google Maps. Lebih jauh, dibandingkan penelitian Dhamma dan Barus (2025) yang membandingkan performa pengklasifikasi konvensional pada ulasan perhotelan, studi ini membuktikan bahwa integrasi *RandomOverSampler* dan analisis konflik rating–teks mampu mengatasi jebakan akurasi tinggi pada data tidak seimbang sekaligus mengungkap keluhan tersembunyi (*hedged praise*) pada ulasan berbintang tinggi. Perspektif manajemen strategis temuan ini memberikan landasan empiris bagi pengelolaan reputasi digital (*Online Reputation Management* / ORM) dan peningkatan *customer experience* sebagaimana ditekankan oleh Mardikaningsih dan Riski (2026) serta Widiarta et al. (2024). Pertama, tingginya sentimen positif pada aspek Tempat (97,41%) perlu dipertahankan sebagai jangkar narasi pemasaran digital melalui konsistensi kebersihan, kenyamanan tata ruang, dan pemeliharaan titik estetika kafe. Kedua, tingginya volume absolut keluhan pada aspek Servis (23 ulasan negatif) menuntut intervensi operasional pada jam sibuk, seperti penerapan sistem nomor antrean digital (*digital queue management*), penataan alur pesanan *takeaway* terpisah dari *dine-in*, pelatihan komunikasi empatik bagi staf dan petugas parkir, serta respons tanggap terhadap ulasan negatif di *Google Reviews* untuk memulihkan kepercayaan pelanggan. Ketiga, tingginya proporsi sentimen negatif pada aspek Harga (43,48%) menunjukkan pentingnya transparansi porsi, konsistensi standar penyajian kopi dan *pastry*, serta penyediaan paket kombo sarapan (*breakfast bundle*) agar nilai yang dirasakan pelanggan sepadan dengan harga yang dibayarkan.

4.2 Saran

Bagi pengelola Café Gigi Susu, disarankan untuk memprioritaskan perbaikan alur layanan pada jam sibuk melalui pengelolaan antrean digital, pemisahan jalur pemesanan *takeaway* dan *dine-in*, peningkatan komunikasi petugas keamanan dan parkir, serta penawaran paket menu sarapan (*breakfast bundle*) guna meningkatkan persepsi *value-for-money* pelanggan. Pengelola juga disarankan

melakukan pemantauan *Google Reviews* secara berkala dan merespons ulasan negatif secara konstruktif sebagai bagian dari strategi *Online Reputation Management*. Bagi penelitian selanjutnya, disarankan untuk mengeksplorasi metode *Aspect-Based Sentiment Analysis* (ABSA) berbasis pembelajaran mendalam seperti *IndoBERT* atau *RoBERTa* guna menekan porsi ulasan tidak terkategori serta meningkatkan nilai *Precision* pada kelas minoritas.

DAFTAR PUSTAKA

Arminda, N. F., Sulistiyowati, N., & Padilah, T. N. (2023). Implementasi Algoritma Multinomial Naive Bayes Pada Analisis Sentimen Terhadap Ulasan Pengguna Aplikasi Brimo. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(3), 1817–1822.

Azizah, H., Syuhada, F. S., & Sa'dati, Y. (2024). Sentimen Analisis Tempat Wisata Berdasrkan ULasan Google Maps Menggunakan Metode Naive Bayes (Studi Kasus Bukit Merese). *SainsTech Innovation Journal*, *7*(2), 467–475.

Balit, M. N. B., & Utomo, F. S. (2024). Sentiment Analysis of pegipegi. com Review on Google Play Store with Naïve Bayes. *Sistemasi: Jurnal Sistem Informasi*, *13*(3), 1044–1053.

Coletta, L. F. S., da Silva, N. F. F., Hruschka, E. R., & Hruschka, E. R. (2014). Combining classification and clustering for tweet sentiment analysis. *2014 Brazilian Conference on Intelligent Systems*, 210–215.

Dhamma, Y. A., & Barus, S. P. (2025). Sentiment Analysis on Google Reviews Using Naive Bayes, K-Nearest Neighbors, and Logistic Regression to Improve Novotel Services. *Journal of Applied Informatics and Computing*, *9*(1), 106–114.

Elistiana, K. M., Kusuma, B. A., Subarkah, P., & Rozaq, H. A. A. (2023). Improvement of Naive Bayes Algorithm in Sentiment Analysis of Shopee Application Reviews on Google Play Store. *Jurnal Teknik Informatika (Jutif)*, *4*(6), 1431–1436.

Fatikhah, T. Z., Muhtarom, M., & Oktaviani, I. (2024). Analisis sentimen ulasan Mie Gacoan Solo Veteran di Google Maps menggunakan algoritma Naive Bayes. *ZONAsi: Jurnal Sistem Informasi*, *6*(3), 649–658.

Firdaus, M. F., Ratnawati, D. E., & Setiawan, N. Y. (2024). Analisis Sentimen Berbasis Aspek Pada Ulasan Pelanggan Restoran Menggunakan Algoritma Support Vector Machine (Studi Kasus: Depot Bamara). *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *11*(6), 1265–1272.

Hizham, F. A., Asy'ari, H., & Urrochman, M. Y. (2024). Sentiment Analysis Of Ijen Crater Reviews Using Naïve Bayes Classification And Oversampling Optimization. *Sistemasi: Jurnal Sistem Informasi*, *13*(5), 2090–2103.

Inyang, E., & White, J. (2026). From Reviews to Revenue: Examining the Impact of Online Reputation Management on Small Business Performance. *Journal of Small Business Strategy*, *36*(1).

Julianto, Y., Setiabudi, D. H., & Rostianingsih, S. (2022). Analisis Sentimen Ulasan Restoran Menggunakan Metode Support Vector Machine. *Jurnal Infra*, *10*(1), 1–7.

Kamal, W. W., & Ratnasari, C. I. (2021). Analisis Sentimen Ulasan Produk: Kajian Pustaka. *AUTOMATA*, *2*(1).

Mardikaningsih, R., & Riski, M. (2026). Daya Saing UMKM di ShopeeFood: Bagaimana Rating dan Review Mempengaruhi Keputusan Pembelian pada Restoran Non-Franchise. *Jurnal Mahasiswa Manajemen Dan Akuntansi*, *5*(1), 657–676.

Pebdika, A., Herdiana, R., & Solihudin, D. (2023). Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima Pip. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(1), 452–458.

RIZKI, A. F., & Prihartono, W. (2025). Analisis sentimen ulasan Google Maps Rumah Sakit Khalishah di Cirebon dengan algoritma Naïve Bayes. *Jurnal Informatika Dan Teknik Elektro Terapan*, *13*(2).

Setiawan, B. (2025). A review of sentiment analysis applications in Indonesia between 2023-2024. *Journal Information Engineering and Educational Technology) ISSN*, *2549*, 869X.

Sri, L., & Saiful, N. B. (2024). Analisis Sentimen Berdasarkan Hasil Review Lokasi Google Map Menggunakan Natural Language Toolkit TextBlob dan Naïve Bayes. *JAMI: Jurnal Ahli Muda Indonesia*, *5*(2), 114–126.

Sudiatmika, I. P. G. A., Saputra, P. S., & Rahardian, R. L. (2025). Sentiment analysis of tourist reviews on google maps for pura besakih using machine learning algorithms. *Jurnal Mandiri IT*, *14*(1), 149–158.

Widiarta, I. K. A., Atmaja, N. P. C. D., & Manek, D. (2024). Pengaruh online customer review, kewajaran harga, dan kualitas pelayanan terhadap keputusan menginap di Hotel Holiday Inn Express Baruna Bali. *Emas*, *5*(10), 126–145.