

**TUGAS AKHIR - ES234849**

**ANALISIS SENTIMEN BERBASIS ASPEK DAN  
KATEGORISASI ASPEK PADA TEKS BAHASA INGGRIS  
DENGAN TEKNIK TRANSFER LEARNING (STUDI  
KASUS : ULASAN HOTEL DI PROVINSI BALI)**

**Chang Hong Kwang**  
NRP 5026201127

Dosen Pembimbing  
**Renny Pradina, S.T.,M.T**  
NIP 1981071720122001

**Program Studi Sarjana**  
Departemen Sistem Informasi  
Fakultas Teknologi Elektro dan Informatika Cerdas  
Institut Teknologi Sepuluh Nopember  
Surabaya  
2024



**TUGAS AKHIR - ES234849**

**ANALISIS SENTIMEN BERBASIS ASPEK DAN  
KATEGORISASI ASPEK PADA TEKS BAHASA INGGRIS  
DENGAN TEKNIK TRANSFER LEARNING (STUDI  
KASUS : ULASAN HOTEL DI PROVINSI BALI)**

**Chang Hong Kwang**  
NRP 5026201127

Dosen Pembimbing  
**Renny Pradina, S.T.,M.T**  
NIP 1981071720122001

**Program Studi Sarjana Sistem Informasi**  
Departemen Sistem Informasi  
Fakultas Teknologi Elektro dan Informatika Cerdas  
Institut Teknologi Sepuluh Nopember  
Surabaya  
2024



**FINAL PROJECT - ES234849**

**ASPECT BASED SENTIMENT ANALYSIS AND ASPECT  
CATEGORIZATION ON ENGLISH TEXT USING  
TRANSFER LEARNING TECHNIQUES (CASE STUDY :  
HOTEL REVIEWS IN BALI PROVINCE)**

**CHANG HONG KWANG**  
NRP 5026201127

Advisor  
**Renny Pradina, S.T.,M.T**  
NIP 1981071720122001

**Information System Undergraduate**  
Department of Information System  
Faculty of Intelligent Electrical and Informatics Technology  
Institut Teknologi Sepuluh Nopember  
Surabaya  
2024

# LEMBAR PENGESAHAN

**ANALISIS SENTIMEN BERBASIS ASPEK DAN EKSTRAKSI KATEGORI  
PADA TEKS BAHASA INGGRIS DENGAN TEKNIK TRANSFER LEARNING  
(STUDI KASUS : ULASAN HOTEL DI PROVINSI BALI)**

## TUGAS AKHIR

Disusun Untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Komputer (S.Kom)  
pada

Departemen Sistem Informasi  
Fakultas Teknologi Elektro dan Informatika Cerdas (ELECTICS)  
Institut Teknologi Sepuluh Nopember

Oleh :

**Chang Hong Kwang**

**NRP: 5026201127**

Surabaya, 01 Agustus 2024

**Kepala Departemen Sistem Informasi**

**LP/P/24/163**

**DSI-130-F-23-267/9**



**Dr. Mudjahidin, ST, MT**

**NIP. 197010102003121001**

**LEMBAR PERSETUJUAN****ANALISIS SENTIMEN BERBASIS ASPEK DAN EKSTRAKSI KATEGORI  
PADA TEKS BAHASA INGGRIS DENGAN TEKNIK TRANSFER LEARNING  
(STUDI KASUS : ULASAN HOTEL DI PROVINSI BALI)****TUGAS AKHIR**

Disusun Untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Komputer  
pada

Departemen Sistem Informasi  
Fakultas Teknologi Elektro dan Informatika Cerdas  
Institut Teknologi Sepuluh Nopember

Oleh :

**Chang Hong Kwang**


**NRP: 5026201127**

Disetujui Tim Penguji:


Tanggal Ujian:  
Periode Wisuda:

24 Juli 2024  
September 2024

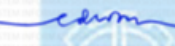
**Renny Pradina, S.T, M.T**

  
(Pembimbing 1)

**Faizal Mahananto, S.Kom, M.Eng, Ph.D**

  
(Penguji 1)

**Edwin Riksakomara, S.Kom, M.T**

  
(Penguji 2)

## PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini:

Nama mahasiswa / NRP : Chang Hong Kwang / 5026201127  
Program studi : Sarjana  
Dosen Pembimbing / NIP : Renny Pradina, S.T, M.T / 198107172012122001

dengan ini menyatakan bahwa Tugas Akhir dengan judul "ANALISIS SENTIMEN BERBASIS ASPEK DAN EKSTRAKSI KATEGORI PADA TEKS BAHASA INGGRIS DENGAN TEKNIK TRANSFER LEARNING (STUDI KASUS : ULASAN HOTEL DI PROVINSI BALI)" adalah hasil karya sendiri, bersifat orisinal, dan ditulis dengan mengikuti kaidah penulisan ilmiah.

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia menerima sanksi sesuai dengan ketentuan yang berlaku di Institut Teknologi Sepuluh Nopember.

Surabaya, 31 Juli 2024

Mengetahui

Dosen Pembimbing

Mahasiswa



Renny Pradina, S.T, M.T  
NIP. 198107172012122001

Chang Hong Kwang  
NRP. 5026201127

## ABSTRAK

**Nama Mahasiswa / NRP** : Chang Hong Kwang/ 5026201127  
**Departemen** : Sistem Informasi  
**Dosen Pembimbing** : Renny Pradina, S.T., M.T.

### Abstrak

Pariwisata merupakan sektor yang sedang dikembangkan oleh pemerintah karena mempunyai peran vital terhadap pembangunan nasional dan pendapatan negara. Menurut data dari Kemenparekraf, kunjungan wisatawan mancanegara ke Indonesia pada Januari-Juli 2023 meningkat sebesar 196,85% dari bulan Januari-Juli 2022. Ini merupakan indikator besarnya potensi pariwisata di Indonesia.

Akomodasi merupakan hal yang penting dalam sektor pariwisata. Seringkali wisatawan akan mencari akomodasi yang cocok sesuai preferensi masing masing berdasarkan teks ulasan di platform-platform seperti Google Maps, Traveloka, Trip Advisor, PegiPegi. Ulasan-ulasan yang banyak ini menjadi kesulitan tersendiri untuk mengelola umpan balik wisatawan. Hal ini dapat diselesaikan dengan analisis sentimen berbasis aspek. Akan tetapi, pada analisis berbasis aspek sering kali akan muncul aspek yang sangat banyak dan mengakibatkan kesulitan menganalisa data. Oleh karena itu, pada task ABSA, dikenalkan ASQE (Aspect Sentiment Quadruple Extraction) yang mengekstrak 4 elemen yaitu aspek, opini, kategori aspek, dan polaritas dalam satu kalimat. Penelitian ini akan menghasilkan model ASQE yang bermanfaat untuk menganalisis teks ulasan berbahasa Inggris hotel di Bali.

Berdasarkan hasil penelitian, model pre-trained berbasis T5 mampu memiliki performa yang cukup baik untuk melakukan ekstraksi *quadruple* pada domain teks berbahasa Inggris. Performa yang dicapai model yaitu mencapai nilai F1 sebesar 0.557 yang mendekati model baseline berdasarkan referensi. Secara kuantitatif, model dapat mengekstrak *quadruple* dengan cukup baik dan model dapat memahami teks dengan baik. Selain itu, model juga terbukti dapat diimplementasikan pada teks ulasan yang banyak dan menarik *insight* yang komprehensif ketimbang membaca teks ulasan secara manual, seperti melihat perbandingan sentimen secara keseluruhan, sebaran aspek setiap kategori, perbandingan sentimen setiap kategori, perbandingan sentimen setiap sub-kategori, serta analisis kelebihan dan kekurangan hotel.

**Kata kunci:** ASQE, pariwisata, sentimen analisis berbasis aspek,

## ABSTRACT

**Student Name / NRP** : Chang Hong Kwang/ 5026201127  
**Department** : Informastion System  
**Advisor** : Renny Pradina, S.T., M.T.

### *Abstract*

*The government is presently focusing on the development of the tourism sector due to its crucial role in national progress and government revenue. As per information from the Ministry of Tourism and Creative Economy (Kemenparekraf), the number of foreign tourists visiting Indonesia between January and July 2023 has surged by 196.85% in comparison to the corresponding period in 2022, indicating the substantial potential of the tourism industry in Indonesia.*

*A fundamental aspect of the tourism sector is accommodation. Travelers frequently look for appropriate lodging based on their preferences, relying on reviews from platforms like Google Maps, Traveloka, Trip Advisor, and PegiPegi. The sheer volume of these reviews presents a challenge in effectively managing tourist feedback. To address this, aspect-based sentiment analysis is employed. However, such analysis often results in a multitude of aspects, complicating the data analysis process. Consequently, in the Aspect-Based Sentiment Analysis (ABSA) task, ASQE (Aspect Sentiment Quadruple Extraction) is introduced to extract four key elements – aspect, opinion, aspect category, and polarity – from a single sentence. Thus, this study will utilize the ASQE model to evaluate English-language hotel reviews in Bali, generating four distinct components*

*Based on the research results, the pre-trained model based on T5 demonstrated quite good performance in extracting quadruples in the domain of English text. The model achieved an F1 score of 0.557, which is close to the baseline model according to the reference. Quantitatively, the model can extract quadruples well and understand the text effectively. Additionally, the model has been proven to be implementable on large review texts, drawing comprehensive insights compared to manually reading review texts, such as overall sentiment comparison, aspect distribution in each category, sentiment comparison in each category, sentiment comparison in each sub-category, as well as analysis of the strengths and weaknesses of hotels*

**Keywords:** *ASQE, tourism, aspect based sentiment analysis*



## KATA PENGANTAR

Segala puji dan rasa syukur penulis panjatkan kepada Tuhan Yang Maha Esa, Allah SWT, atas seluruh rahmat-Nya, sehingga penelitian tugas akhir dengan judul "Analisis Sentimen Berbasis Aspek dan Kategorisasi Aspek pada Teks Bahasa Inggris dengan teknik Transfer Learning (studi kasus: teks ulasan hotel di Bali)" dapat diselesaikan dengan baik dan tepat waktu. Laporan ini dibuat untuk memenuhi persyaratan kelulusan Program Sarjana S1 Departemen Sistem Informasi, Fakultas Teknologi Elektro dan Informatika Cerdas (FTEIC).

Penulis mengucapkan terima kasih kepada semua pihak yang telah membantu dan mendukung penyelesaian tugas akhir ini, baik dalam bentuk dukungan materiil, moril, saran, dan doa sehingga penelitian ini dapat berlangsung dengan lancar. Secara khusus, penulis mengucapkan terima kasih kepada:

1. Keluarga dan teman-teman yang selalu memberikan dukungan, doa, dan motivasi selama menjalani perkuliahan hingga saat ini.
2. Ibu Renny Pradina, S.T., M.T., sebagai dosen pembimbing yang telah meluangkan waktu dan tenaga untuk membimbing, memberikan konsultasi, serta dukungan moral dan materiil selama pembuatan tugas akhir ini.
3. Bapak Faizal Mahananto, S.Kom., M.Eng., Ph.D., dan Edwin Riksakomara, S.Kom., M.T., sebagai dosen penguji yang telah memberikan kritik dan saran untuk perbaikan tugas akhir ini.
4. Bapak Faizal Mahananto, S.Kom., M.Eng., Ph.D., juga sebagai dosen wali yang telah memberikan bantuan konsultasi serta dukungan akademis selama menjalani perkuliahan.
5. Johana Gracia, yang telah membantu dan meluangkan waktunya dalam untuk berdiskusi dan membantu baik di dalam maupun di luar pengerjaan tugas akhir serta dukungan moralnya yang sangat besar.

Penulis menyadari bahwa tugas akhir ini masih jauh dari sempurna dan terbuka untuk perbaikan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun. Semoga tugas akhir ini dapat bermanfaat bagi para pembaca.

Surabaya, 23 Juli 20

Chang Hong Kwang

## DAFTAR ISI

BAB 1	PENDAHULUAN .....	1
1.1	Latar Belakang.....	1
1.2	Rumusan Masalah.....	3
1.3	Batasan Masalah .....	3
1.4	Tujuan.....	3
1.5	Manfaat.....	3
1.5.1	Manfaat Praktis.....	3
1.5.2	Manfaat Teoritis .....	3
1.6	Relevansi .....	3
BAB 2	TINJAUAN PUSTAKA.....	5
2.1	Hasil Penelitian Terdahulu .....	5
2.2	Dasar Teori .....	6
2.2.1	Pengolahan Bahasa Alami.....	6
2.2.2	Analisis Sentimen Berbasis Aspek.....	6
2.2.3	Transformers.....	7
2.2.4	Text-to-Text Transfer Transformer (T5).....	8
2.2.5	ASQE(Asspect Sentiment Quadruple Extraction) .....	9
2.2.6	Metrik Evaluasi.....	10
BAB 3	METODOLOGI .....	11
3.1	Alur Pengerjaan Tugas Akhir .....	11
3.2	Urutan pelaksanaan penelitian.....	12
3.2.1	Identifikasi Masalah .....	12
3.2.2	Pra-pemrosesan data .....	12
3.2.3	Pembuatan Model .....	12
3.2.4	Analisis Sentimen .....	13
3.2.5	Penulisan Tugas Akhir.....	14
BAB 4	HASIL DAN PEMBAHASAN .....	15
4.1	Pra-pemrosesan data .....	15
4.2	Skenario dataset.....	15
4.3	Pelabelan data .....	16
4.4	Eksplorasi data.....	17
4.5	Training Data .....	26
4.5.1	Model Baseline .....	26
4.5.2	Persiapan dataset.....	26
4.5.3	Metode Training .....	27

4.5.4	Hasil Model .....	27
4.5.5	Optimasi.....	31
4.5.6	Pengkategorian .....	31
4.5.7	Perbandingan antar model .....	32
4.6	Implementasi Model dan Penggunaan Kategorisasi.....	33
4.6.1	Dataset .....	33
4.6.2	Pre-processing .....	33
4.6.3	Analisis data .....	33
BAB 5	Kesimpulan dan Saran .....	49
5.1	Kesimpulan.....	49
5.2	Saran .....	49
	DAFTAR PUSTAKA .....	51
	LAMPIRAN.....	54
	Lampiran 1: Contoh Kelas ATE dan APC .....	54
	Lampiran 2: Link Github Repository .....	55
	BIODATA PENULIS .....	56

## DAFTAR GAMBAR

Gambar 2.1 Arsitektur Model Transformers (sumber : (Vaswani et al., 2017))	7
Gambar 2.2 T5	8
Gambar 2.3 ASQE	9
Gambar 3.1 Bahan dan Peralatan yang Digunakan	11
Gambar 4.1 Distribusi Sentimen	18
Gambar 4.2 Distribusi word length single sentence dataset	18
Gambar 4.3 Distribusi word length multi sentences dataset	19
Gambar 4.4 Distribusi quadruple multi sentences dataset	20
Gambar 4.5 Distribusi quadruple single sentence dataset	20
Gambar 4.6 Distribusi Jarak Aspek dan Opini multi sentences dataset	21
Gambar 4.7 Distribusi Jarak Aspek dan Opini multi sentences dataset	21
Gambar 4.8 Visualisasi Distribusi Kategori	22
Gambar 4.9 Visualisasi Distribusi Sub-Kategori Lokasi	22
Gambar 4.10 Visualisasi Distribusi Sub-Kategori Hotel	23
Gambar 4.11 Visualisasi Distribusi Sub-Kategori "FOOD"	24
Gambar 4.12 Visualisasi Distribusi Sub-Kategori "FACILITY"	24
Gambar 4.13 Visualisasi Distribusi Sub-Kategori "FACILITY"	25
Gambar 4.14	25
Gambar 4.15 Pie Chart Distribusi Polaritas	34
Gambar 4.16 Bar Chart Distribusi Polaritas	34
Gambar 4.17 Sebaran Aspek Kategori Layanan	35
Gambar 4.18 Sebaran Aspek Kategori Makanan	35
Gambar 4.19 Sebaran Aspek Kategori Hotel	36
Gambar 4.20 Sebaran Aspek Kategori	36
Gambar 4.21 Sebaran Aspek Lokasi	37
Gambar 4.22 Sebaran Aspek Kamar	37
Gambar 4.23 Sebaran Aspek Fasilitas	38
Gambar 4.24 Distribusi Sentimen Kategori	39
Gambar 4.25 Distribusi Sentimen Positif	40
Gambar 4.26 Distribusi Sentimen Negatif	40
Gambar 4.27 Distribusi Polaritas Kedua	41
Gambar 4.28 Distribusi Polaritas Fasilitas	41
Gambar 4.29 Distribusi Polaritas Kategori Lokasi	41
Gambar 4.30 Distribusi Polaritas Kategori Makanan	41
Gambar 4.31 Distribusi Polaritas Kategori Layanan	41
Gambar 4.32 Distribusi Polaritas Kategori Hotel	41
Gambar 4.33 Distribusi Polaritas Kategori Kamar	42
Gambar 4.34 Distribusi Sub Kategori Lokasi	42
Gambar 4.35 Wordcloud opini lokasi	43
Gambar 4.36 wordcloud Kategori "SERVICE"	44
Gambar 4.37 Bar-chart sub-kategori "ROOM"	45
Gambar 4.38 Wordcloud kategori "ROOM"	46
Gambar 4.39 Bar chart Distribusi Kategori "FOOD"	47
Gambar 4.40	47

## DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu.....	5
Tabel 4.1 Contoh dataset .....	15
Tabel 4.2 Daftar Kategori bagian 1 .....	16
Tabel 4.3 Daftar Kategori bagian 2 .....	16
Tabel 4.4 Jumlah Kalimat dan quadruple dataset.....	17
Tabel 4.5 Baseline Model.....	26
Tabel 4.6 Hasil Training.....	28
Tabel 4.7 Contoh Prediksi Model SS .....	30
Tabel 4.8 Hasil Model SS setelah optimasi kategori.....	31
Tabel 4.9 Performa model SS per task .....	31
Tabel 4.10 Performa model pada kategori utama.....	32
Tabel 4.11 Performa model pada sub-kategori.....	32
Tabel 4.12 Perbandingan contoh prediksi model ss dan model ms.....	32
Tabel 4.13 Contoh implementasi prediksi model machine learning .....	33
Tabel 4.14 Jumlah Sentimen Positif, Negatif, dan Netral .....	34
Tabel 4.15 Contoh quadruple yang membahas lokasi.....	43
Tabel 4.16 Contoh teks yang membahas layanan.....	44
Tabel 4.17 Contoh Teks yang Membahas “ROOM”.....	46
Tabel 4.18 Contoh Teks yang Membahas “FOOD” .....	48

## DAFTAR PERSAMAAN

Persamaan 2.1 Rumus F1 .....	9
Persamaan 2.2 Rumus Akurasi.....	9
Persamaan 2.3.Rumus Presisi.....	10
Persamaan 2.4.Rumus <i>Recall</i> .....	10

# BAB 1 PENDAHULUAN

## 1.1 Latar Belakang

Saat ini pariwisata termasuk ke dalam sektor yang sedang dikembangkan oleh pemerintah Indonesia karena sektor ini mempunyai peran yang vital terhadap pembangunan nasional dan berperan dalam pendapatan daerah ataupun negara (Aliansyah & Hermawan, 2021). Tidak hanya itu, sektor pariwisata juga dianggap mampu mengurangi angka pengangguran (Aliansyah & Hermawan, 2021). Saat ini, angka kunjungan wisatawan mancanegara meningkat drastis. Pada bulan Januari – Juli 2023, jumlah kunjungan kumulatif sebesar 6.312.387 kunjungan yang meningkat sebesar 196,85% dari bulan Januari – Juli 2022, yaitu sebesar 2.126.436 kunjungan (Pusdatin Kemenparekraf, 2023). Ini salah satu indikator yang menunjukkan betapa besarnya potensi pariwisata Indonesia. Salah satu bagian dari pariwisata adalah akomodasi. Akomodasi sering kali menjadi pengeluaran terbesar wisatawan saat berwisata di Indonesia, sehingga akomodasi merupakan bagian yang detrimental dalam sektor pariwisata (Badan Pusat Statistik, 2021).

Sering kali saat melakukan wisata, para wisatawan akan mencari akomodasi yang cocok bagi mereka. Teks ulasan sering digunakan untuk menjadi patokan bagi *customer* terkait informasi produk dan membantu *customer* dalam membuat keputusan (Liu, 2012). Saat ini, para wisatawan dapat melihat ulasan-ulasan tempat akomodasi di beberapa platform seperti Google Maps, Traveloka, TripAdvisor, PegiPegi, dan lain lain. Ulasan yang sangat banyak akan menjadi kesulitan tersendiri untuk mengelola umpan balik dari wisatawan (Muhammad et al., 2021). Tidak hanya itu, sering kali ulasan maupun *rating* yang ada pada ulasan hotel di *platform* tersebut terkadang kurang membantu bagi sebagian orang. Sebagai contoh, orang yang mementingkan makanan hotel yang enak kemungkinan akan kesusahan dalam mencari hotel dengan kondisi tersebut dan membutuhkan waktu yang ekstra untuk memfilter hotel dengan syarat tersebut. Hal ini dapat diselesaikan dengan analisis sentimen. Analisis sentimen adalah proses mengolah data teks dari bentuk tidak terstruktur menjadi informasi sentimen yang bernilai positif maupun negatif (Liu, 2012). Dengan analisis sentimen, ulasan yang banyak dapat diolah dan dianalisis. Analisis sentimen secara konvensional menentukan sentimen keseluruhan suatu teks, baik itu positif, negatif, maupun netral.

Dalam suatu teks, penulis dapat menyatakan beberapa sentimen sekaligus. Sebagai contoh, dalam kalimat “Toiletnya kotor, tetapi makanannya enak sekali.” Di kalimat ini muncul sentimen positif terhadap makanan dan sentimen negatif dalam aspek toilet. Analisis sentimen tradisional tidak dapat memprediksi secara akurat polaritas untuk aspek spesifik karena analisis sentimen tradisional tidak mempertimbangkan polaritas yang lebih rinci dari berbagai aspek (Zeng et al., 2019). Aspect Based Sentiment Analysis (ABSA) melibatkan prediksi polaritas sentimen independen untuk aspek yang ditargetkan dalam kalimat atau bacaan yang sama, tidak seperti analisis sentimen tradisional yang menyimpulkan polaritas sentimen secara umum dan menyeluruh (Zeng et al., 2019).

Penelitian-penelitian sebelumnya telah menggunakan ABSA untuk menganalisis sentiment, seperti penelitian yang dilakukan oleh (Fang & Tao, 2019). Penelitian ini menggunakan studi kasus ulasan konsumen online dengan menggunakan metode *transfer learning* dari BERT. Di penelitian ini, terbukti penggunaan *transfer learning* pada ABSA memiliki akurasi yang lebih tinggi ketimbang approach lain seperti *label powerset*, *binary relevance*, dan *algorithm adaptation*. Selain itu, penelitian yang dilakukan oleh (Kusuma, 2023) menunjukkan

pengembangan model sentiment analisis berbasis aspek menggunakan teknik *transfer learning* dan *local sentiment aggregation* dengan studi kasus ulasan pariwisata di provinsi Bali. Penelitian ini menunjukkan dengan *local sentiment aggregation* (LSA) dapat menghasilkan performa yang baik pada data Bahasa Inggris dan Bahasa Indonesia.

Meskipun begitu, analisis sentimen yang berbasis aspek pada ulasan yang sangat banyak dapat memiliki kekurangan, seperti muncul aspek yang sangat banyak dan mengakibatkan kesulitan dalam menganalisis data. Contoh, pada kalimat ulasan “Udangnya bau amis dan tidak segar” dan “ayam gorengnya empuk” pada analisis sentimen berbasis aspek konvensional akan mengekstrak aspek “udang” dan “ayam”, padahal kedua aspek tersebut dapat dikategorikan ke dalam satu aspek yaitu “makanan”. Dengan mengkategorikan aspek aspek yang beragam, tentu dapat mempermudah *stakeholder* untuk menganalisis dan membaca data. Dalam *task* ABSA, terdapat salah satu subtask yakni *Aspect Sentiment Quadruple Extraction (ASQE)*, yang mengekstrak 4 elemen, yaitu aspek, opini, kategori, dan polaritas dalam satu kalimat (Zhang et al., 2022). Sebagai contoh, dalam kalimat “Udangnya bau amis dan tidak segar tetapi ayam gorengnya empuk.” akan mengekstrak quadruple {(udang, bau amis,makanan,negatif),(ayam goreng,empuk,makanan,positif)} dengan urutan quadruple aspek, opini, kategori, dan polaritas. Dengan menggunakan metode ini, deteksi kategori dapat membantu bisnis dalam mengevaluasi kualitas dari berbagai macam aspek barang, yang dapat membantu meningkatkan daya saing pasar (Zhang et al., 2022).

Penelitian yang dilakukan oleh (Ziaulhaq, 2023) menunjukkan penggunaan model (Span-ASTE) untuk menganalisis ulasan berbahasa Inggris tentang hotel di Bali yang menghasilkan aspek, opini, sentimen. Selanjutnya dari dari aspek aspek ini dilakukan kategorisasi dengan menggunakan *Wu-Palmer Similarity Score*. Di dalam penelitian tersebut, kategorisasi aspek menggunakan tingkat kemiripan antar kata aspek. Kelemahan metode ini yaitu beberapa kata yang memiliki arti yang sama dianggap memiliki arti yang berbeda dan tidak dapat mengetahui konteksnya pada ulasan sehingga perlu dilakukan *pre-processing* data secara ekstra. Oleh karena itu, dalam penelitian ini diharapkan dapat dilakukan ekstraksi aspek, opini, polaritas, dan kategorisasi aspek dengan teknik *machine learning*. Penelitian ini juga membandingkan model dari yang dihasilkan dari penelitian (Ziaulhaq, 2023). Oleh sebab itu, penelitian ini menggunakan data yang sama dari penelitian (Ziaulhaq, 2023) yaitu dengan data ulasan 48 hotel dari situs web TripAdvisor yang telah di-*scrap*, dilakukan *preprocessing* dan dilabeli oleh (Ziaulhaq, 2023).

Output dari penelitian ini adalah model machine learning yang dapat mengekstraksi aspek, opini, kategori, dan polaritas dari teks ulasan hotel. Harapannya adalah ketimbang menganalisis teks ulasan dengan membaca satu per satu ulasan secara manual, model machine learning ini dapat membantu untuk menganalisis teks ulasan yang banyak secara keseluruhan. Sebagai contoh, bagi pengelola hotel yang ingin mengetahui apa kekurangan hotelnya dalam periode 3 bulan terakhir, dengan menggunakan model ASQE ini, dapat mengetahui aspek apa yang banyak dinilai kurang dan bagus oleh customer, contohnya kategori aspek kebersihan ternyata memiliki 70% sentimen negatif dan kategori aspek makanan memiliki 80% sentimen positif. Dari segi customer, contohnya customer yang mengutamakan kebersihan hotel, dengan bantuan model ini yang dapat diaplikasikan pada teks ulasan dari berbagai macam hotel, customer dapat memfilter mana hotel yang memiliki jumlah sentimen positif yang banyak terhadap aspek kebersihan.

Pada tugas akhir ini, model *machine learning* yang mengimplementasikan ASQE akan digunakan pada data ulasan hotel di Bali. Bali merupakan provinsi dengan destinasi pariwisata



yang sangat banyak, mulai dari pantai, gunung, sampai acara kelas dunia seperti konser musik dan *event* lari marathon seperti Maybank Bali Marathon. Pada tahun 2022, Bali tercatat sebagai salah satu dari lima provinsi dengan jumlah akomodasi dan hotel terbanyak, yakni sebanyak 3.528 usaha atau sebesar 11,86% persen dari total usaha hotel dan akomodasi di Indonesia. (BPS, 2022). Penelitian ini dilakukan untuk melakukan analisis sentimen berbasis aspek dan kategorisasi aspek yang diharapkan dapat membantu para *stakeholder* untuk meningkatkan kualitas akomodasi, promosi pariwisata, dan lain-lain. Selain itu, dari segi konsumen, penelitian ini diharapkan dapat membantu wisatawan dalam memilih akomodasi yang cocok sesuai preferensi masing masing.

## 1.2 Rumusan Masalah

Rumusan masalah pada tugas akhir ini adalah sebagai berikut :

1. Bagaimana membangun model ASQE untuk mengekstraksi quadruple sentimen aspek?
2. Bagaimana performa model quadruple yang dihasilkan pada ulasan hotel?

## 1.3 Batasan Masalah

Batasan masalah pada tugas akhir ini adalah sebagai berikut :

1. Data yang digunakan adalah ulasan hotel di wilayah Bali
2. Data ulasan diambil dalam bahasa Inggris
3. Model yang di-*train* tidak dilakukan *fine-tuning*
4. Tidak digunakan otomatisasi dalam pembuatan pipeline
5. Tidak dilakukan deployment model

## 1.4 Tujuan

Tujuan dari penelitian ini adalah sebagai berikut :

1. Membangun model ASQE dengan dataset ulasan hotel.
2. Menguraikan hasil quadruple ulasan hotel hasil model ASQE
3. Mengoptimalkan model machine learning
4. Mempraktekkan penggunaan model ASQE pada ulasan hotel

## 1.5 Manfaat

### 1.5.1 Manfaat Praktis

Penelitian tugas akhir ini diharapkan dapat membantu para wisatawan dalam memilih hotel yang sesuai dengan preferensi masing masing. Selain itu, dari sisi pemilik dan pengelola usaha akomodasi dapat meningkatkan pelayanan dan bagian bagian lain sehingga usaha dapat berkembang lebih baik.

### 1.5.2 Manfaat Teoritis

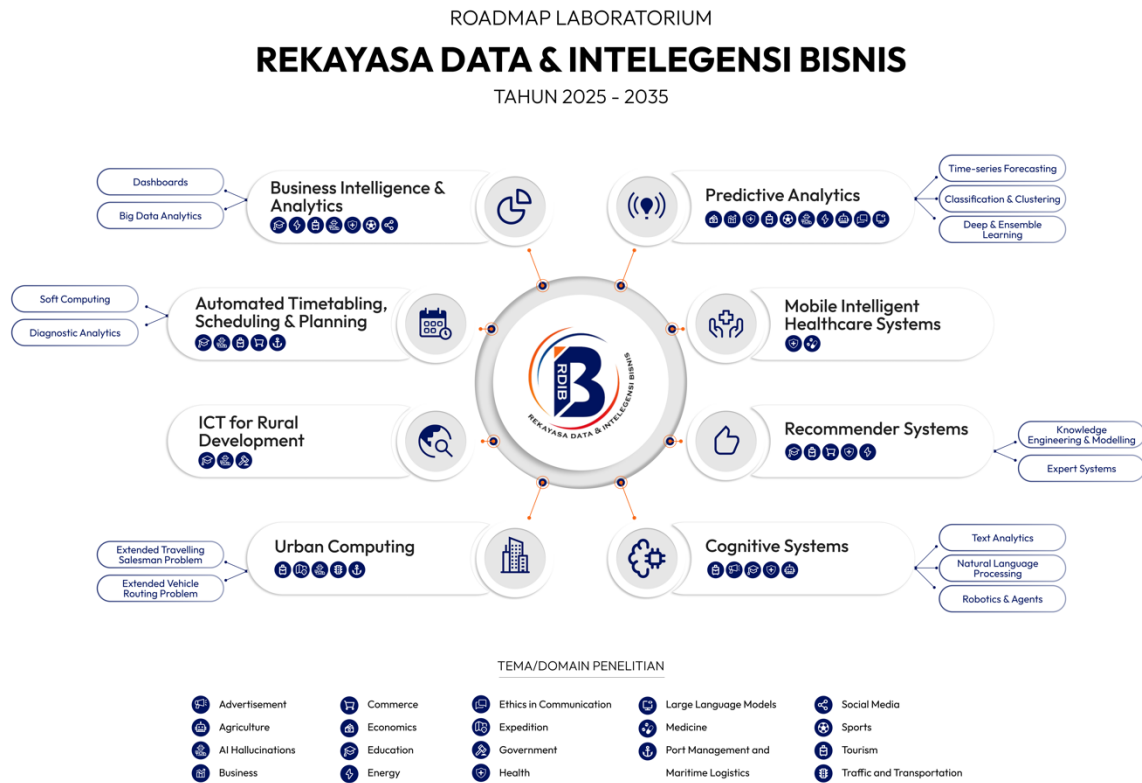
Penelitian tugas akhir ini diharapkan dapat memberi manfaat bagi perkembangan di bidang *Nature Language Processing*. Tidak hanya itu, penelitian ini diharapkan juga dapat menjadi referensi penelitian dan pengembangan aplikasi yang dapat menggunakan implementasi NLP, terutama pada bidang analisis berbasis aspek (ABSA).

## 1.6 Relevansi

Relevansi tugas akhir ini terhadap laboratorium Rekayasa Data dan Intelegensi Bisnis (RDIB) karena tugas ini mengimplementasikan mata kuliah yang beradiah di bawah naungan RDIB yaitu :

- a) Analitika Bisnis
- b) Statistika
- c) Kecerdasan bisnis

Topik pada penelitian ini merupakan analisis sentimen berbasis aspek sehingga untuk relevansi dalam laboratorium Rekayasa Data dan Intelegensi Bisnis (RDIB) terkait dengan tugas akhir ini adalah pemrosesan bahasa alami.



Gambar 1.1 Roadmap Laboratorium Rekayasa & Intelegensi Bisnis (RDIB)

Gambar 1.1 memaparkan *roadmap* untuk laboratorium RDIB, dengan domain penelitian yang berhubungan dengan penelitian ini adalah *cognitive systems*, khususnya di bidang *natural language processing (NLP)*.

## BAB 2 TINJAUAN PUSTAKA

Bab ini berisi tentang penelitian-penelitian terdahulu yang berhubungan dengan tugas akhir dan teori-teori yang melandasi tugas akhir ini.

### 2.1 Hasil Penelitian Terdahulu

Subbab ini berisi tentang referensi penelitian yang berkaitan dengan tugas akhir ini.

Terdapat beberapa rujukan penelitian terdahulu yang berkorelasi dengan penelitian ini yang dapat dijadikan referensi dan pembandingan.

*Tabel 2.1 Penelitian Terdahulu*

<b>Judul</b>	<b>Interactive Attention Networks for Aspect-Level Sentiment Classification</b>
<b>Penulis,Tahun</b>	(Ma et al., 2017)
<b>Metode</b>	Interactive Attention Networks
<b>Gambaran Umum</b>	Interactive Attention Networks terdiri dari dua bagian yang memodelkan target dan konteks secara interaktif. Dengan menggunakan representasi kata sebagai input, peneliti menggunakan jaringan LSTM untuk mendapatkan status tersembunyi kata-kata pada tingkat kata untuk target dan konteks masing-masing. Peneliti mengambil nilai rata-rata dari status tersembunyi target dan konteks untuk mengawasi pembuatan vektor perhatian yang digunakan untuk menangkap informasi penting dalam konteks dan target. Dengan desain ini, target dan konteks saling mempengaruhi pembentukan representasi mereka. Akhirnya, representasi target dan representasi konteks digabungkan sebagai representasi akhir yang dimasukkan ke dalam fungsi softmax untuk klasifikasi sentimen pada tingkat aspek.
<b>Keterkaitan</b>	Implementasi Aspect Based Sentiment Analysis
<b>Judul</b>	A multi-task learning model for Chinese-oriented aspect polarity classification and aspect term extraction
<b>Penulis,Tahun</b>	(Yang et al., 2021)
<b>Metode</b>	<ul style="list-style-type: none"> <li>• Local Context Focus (LCF)</li> <li>• Aspect Term Extraction (ATE)</li> <li>• Aspect Polarity Classification (APC)</li> </ul>
<b>Gambaran Umum</b>	Penelitian ini mengusulkan penggunaan multi task learning dalam melakukan ekstraksi istilah aspek dan klasifikasi polaritas aspek terhadap data teks. Model yang dikembangkan berbasis bahasa Cina dan dapat diaplikasikan ke dalam struktur bahasa Inggris. Model ini dapat bekerja secara bersamaan baik dalam bahasa Inggris dan bahasa Cina untuk mengekstrak aspek dan sentimen terhadap data teks. Model ini telah mencapai kinerja yang stabil setelah diujicobakan di ketujuh dataset ABSA.

<b>Keterkaitan</b>	Menggunakan model pre-trained yang dapat mendukung Bahasa multilingual. Model ini dapat dijadikan referensi untuk dicoba saat diimplementasikan ke dataset ulasan hotel.
<b>Judul</b>	ANALISIS DAN KATEGORISASI TRIPLET SENTIMEN LEVEL ASPEK PADA ULASAN HOTEL MENGGUNAKAN SPAN-ASTE DAN WU-PALMER SIMILARITY SCORE (STUDI KASUS: PROVINSI BALI)
<b>Penulis,Tahun</b>	(Ziaulhaq, 2023)
<b>Metode</b>	Span-ASTE Wu-Palmer Similarity Score
<b>Gambaran Umum</b>	Penelitian ini menggunakan model span level <i>Aspect Sentiment Triplet Extraction</i> dengan dataset ulasan hotel di Bali. Dari model ini akan dihasilkan tiga komponen yaitu aspek, opini, dan sentiment. Selanjutnya dari aspek aspek ini akan dikategorisasikan dengan Wu-Palmer Similarity Score
<b>Keterkaitan</b>	Mengekstraksi empat komponen( <i>quadruple</i> ) yaitu aspek, opini, sentiment, dan kategori aspek

## 2.2 Dasar Teori

Subbab ini berisi teori yang berkaitan dengan tugas akhir dan sebagai landasan pengerjaan tugas akhir ini.

### 2.2.1 Pengolahan Bahasa Alami

Pengolahan bahasa alami (NLP) adalah teknologi yang melibatkan machine learning yang mengizinkan computer untuk menafsirkan, memanipulasi, dan mengerti bahasa manusia. Keberadaan NLP sangat penting karena mampu menganalisis secara efisien data teks dan ucapan secara komprehensif. Teknologi ini dapat menavigasi berbagai variasi dialek, bahasa informal, dan variasi tata bahasa yang khas dalam percakapan sehari-hari. Dalam bisnis, NLP digunakan untuk menyederhanakan, mengotomatisasi, dan meningkatkan efisiensi operasi. Contohnya dalam pengolahan dokumen sensitive di sektor asuransi, hukum, dan perawatan kesehatan, serta analitik bisnis untuk mendapatkan pemahaman yang lebih tentang perasaan pelanggan terhadap produk atau layanan seperti dalam mengukur sentimen dan emosi pelanggan. (AWS, 2023)

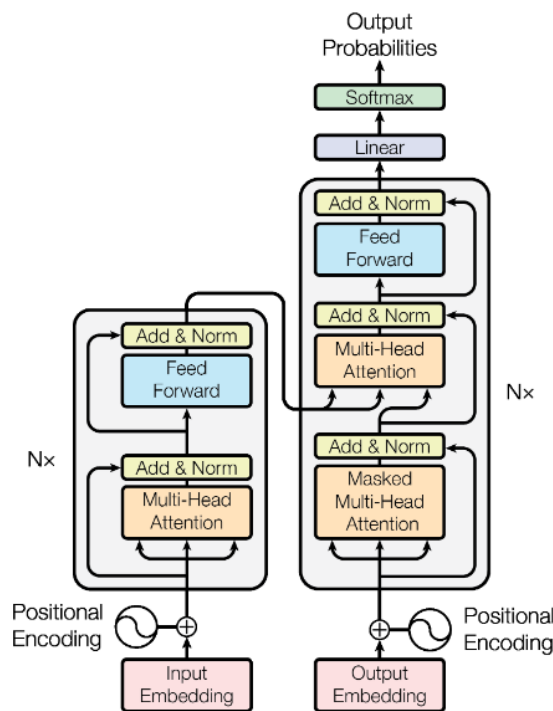
### 2.2.2 Analisis Sentimen Berbasis Aspek

Analisis sentimen merupakan proses eksplorasi data teks yang bertujuan untuk menemukan dan memahami opini, emosi, serta makna yang terkandung di dalamnya (Gupta, 2018). Tujuan utama dari analisis sentimen adalah untuk mengidentifikasi sudut pandang, emosi, atau interpretasi dari kalimat teks atau pidato (Raj, 2021). Hasil dari analisis sentimen ini bisa berupa penanda data yang menunjukkan apakah sentimen tersebut positif, negatif, atau netral. Di sektor bisnis, analisis sentimen membantu dalam memahami pandangan publik terhadap merek, memungkinkan pemangku kepentingan untuk merumuskan strategi respons yang terbaik terhadap sentimen tersebut.

Analisis sentimen tradisional hanya fokus pada polaritas sentimen pada tingkat kalimat atau dokumen secara keseluruhan. Dengan analisis sentimen tradisional, kalimat yang mempunyai lebih dari satu sentimen dalam satu kalimat tidak dapat dideteksi. *Aspect Based Sentiment Analysis* (ABSA) merupakan bagian dari NLP yang memprediksi polaritas sentimen independen untuk aspek yang ditargetkan dalam kalimat atau dokumen yang sama (Zeng et al., 2019). Contohnya pada kalimat “Makanannya enak, tetapi toiletnya kotor” pada ulasan restoran menyatakan bahwa *customer* menyukai makanannya, tetapi tidak dengan toiletnya. Tidak seperti analisis sentimen tradisional, dengan ABSA sentimen dari aspek yang spesifik dapat diekstrak seperti pada aspek makanan kalimat tersebut memiliki sentimen positif dan pada aspek toilet bersentimen negatif. Oleh karena itu, dengan analisis sentimen berbasis aspek, suatu kalimat dapat diekstraksi informasi yang lengkap.

### 2.2.3 Transformers

Transformers dalam NLP adalah arsitektur yang inovatif yang bertujuan untuk menyelesaikan tugas-tugas urutan ke urutan sambil menangani ketergantungan jarak jauh dengan mudah. Model ini sepenuhnya bergantung pada *self attention mechanism* untuk menghitung representasi dari input dan outputnya tanpa menggunakan RNN atau CNN. (Vaswani et al., 2017). Sebagai contoh, dalam kalimat “ Dia menggunakan kunci yang dia temukan untuk membuka pintu”. Hubungan “kunci” dan “membuka” terpisah beberapa kata di dalam kalimat. Dalam arsitektur neural network seperti RNN dan CNN, menangkap *long range dependencies* menjadi sulit. RNN mempunyai kesulitan mempertahankan informasi dalam serangkaian urutan yang panjang karena masalah gradien yang menghilang atau meledak. CNN, meskipun efisien dalam beberapa tugas, sulit untuk menangkap dependensi (ketergantungan) secara efektif pada jarak yang luas dalam sebuah urutan tanpa menggunakan lapisan



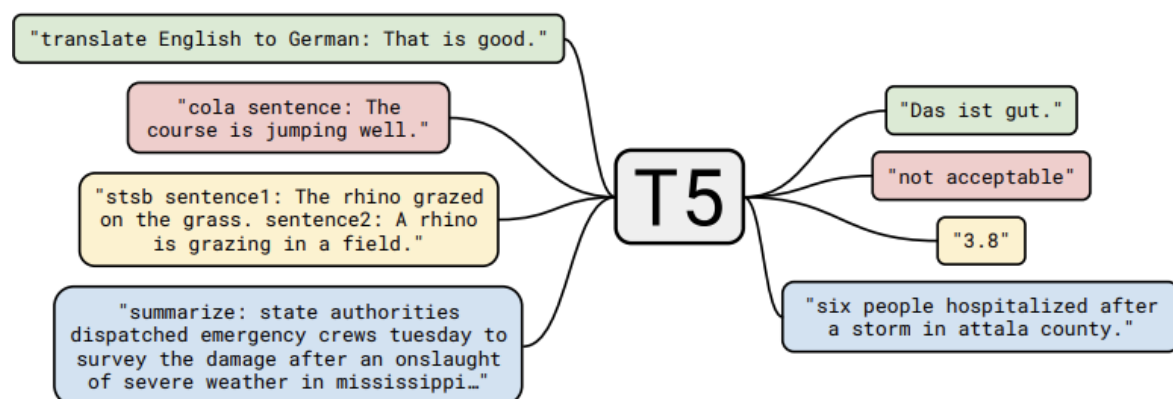
**Gambar 2.1** Arsitektur Model Transformers  
(sumber : (Vaswani et al., 2017))

memanfaatkan mekanisme yang disebut self-attention. Self-attention memungkinkan model untuk menimbang relevansi setiap kata/token dalam urutan terhadap setiap kata/token lainnya. Mekanisme perhatian ini memberdayakan model untuk memberikan tingkatan kepentingan atau relevansi yang berbeda untuk bagian-bagian berbeda dari urutan input, tanpa memandang jaraknya, sehingga memungkinkan pemahaman yang lebih baik terhadap ketergantungan jarak jauh.

Secara ringkas, ketergantungan jarak jauh dalam pemrosesan bahasa alami mewakili hubungan antara kata-kata atau token dalam urutan yang saling terpisah, dan memahami hubungan ini sangat penting untuk memahami konteks dan makna teks. Arsitektur Transformer, melalui mekanisme self-attention-nya, mengatasi ketergantungan jarak jauh ini secara efektif, menjadikannya alat yang kuat dalam tugas-tugas pemrosesan bahasa alami. *ling* yang signifikan. Encoder terletak di area sebelah kiri dan decoder di sebelah kanan. Setiap encoder mempunyai 2 sub layers, yaitu *multi head self attention mechanism* dalam input vektor dan *fully connected feed-forward network* (pahami seperti post processing). Self-attention adalah fungsi yang mengambil tiga vektor  $q$ ,  $k$ , dan  $v$  sebagai input dan menghasilkan vektor  $z$  sebagai output. Vektor  $q$  mewakili pertanyaan yang ingin dijawab oleh model, vektor  $k$  mewakili dokumen yang ingin dijawab, dan vektor  $z$  mewakili jawaban yang dihasilkan oleh model. Transformer menggunakan multi-headed attention, yang menggabungkan beberapa self-attention heads untuk menghasilkan representasi teks yang lebih kompleks. Encoder melakukan mapping terhadap input  $(x_1, \dots, x_n)$  ke *sequence* representasi  $z = (z_1, \dots, z_n)$ . Ini merupakan output dari *self attention* dengan *post-processing*. Dari  $z$ , decoder menghasilkan *output sequence*  $(y_1, \dots, y_m)$ . Setiap decoder memiliki tiga sub layer, yaitu *masked multi-head self attention mechanism* pada output vektor di iterasi sebelumnya, *multi head attention mechanism* pada output dari encoder dan *masked multi headed attention* dari decoder, *fully connected feed forward network* sebagai post processing. Model ini juga menggunakan input dan output preprocessing untuk meningkatkan representasi teks. Input preprocessing melibatkan embedding kata dan menambahkan positional encoding. Output preprocessing melibatkan softmax untuk memilih probabilitas kata terakhir.

## 2.2.4 Text-to-Text Transfer Transformer (T5)

T5 adalah model transformers yang dikenalkan oleh Google Research. Dalam jurnal yang ditulis oleh (Raffel et al., 2019), metode yang dipakai dalam framework T5 adalah dapat memperlakukan semua *task* NLP dalam masalah *text to text* yaitu teks sebagai input dan menghasilkan teks sebagai output. Model T5 di-*train* di *dataset* yang besar yaitu C4 yang berukuran 750 GB teks Bahasa Inggris yang telah dibersihkan. T5 merupakan “text to text”



**Gambar 2.2** Contoh penambahan prefiks sesuai teks pada model T5

framework yaitu semua task NLP seperti *summarization*, *classification*, mempunyai struktur dasar yang sama dengan teks dimasukan ke model sebagai input dan memproduksi teks sebagai output. Karena itu, dengan model yang sama, *loss function* dan *hyperparameter* yang sama dapat digunakan di task-task yang lain seperti yang ditunjukkan pada diagram di bawah.

T5 menggunakan encoder-decoder standar transformers. Meskipun mempunyai arsitektur yang sama dengan Transformers orisinal, T5 dapat mencapai SOTA (State of the Art) karena dua fitur uniknya yaitu *C4 training dataset* dan *text to text framework*. *Text to text* ini dapat dicapai dengan menambahkan prefiks yang spesifik task pada *input sequence* dan *pre-training model* untuk mendapatkan keluaran prefiks yang spesifik. Seperti pada *text summarization task*, dengan *input sequence text* yang panjang pada model lalu menghasilkan teks dengan versi yang ringkas dari input. Untuk itu, perlu menambahkan prefiks “summarize” pada *input sequence*. Pada *language translation task*, input berupa teks dalam suatu bahasa dengan output berupa teks yang sama namun dengan bahasa yang berbeda. Dengan T5, sebagai contoh saat menerjemahkan dari Bahasa Inggris ke Bahasa Jerman maka perlu menambahkan prefiks “translate English to German” pada *input sequence*.

### 2.2.5 ASQE(Aspect Sentiment Quadruple Extraction)

Dalam task ABSA, ada beberapa subtasks seperti Aspect Term Extracntion (ATE), Opinion Term Extraction (OTE), Aspect Polarity Classification (APC), dan lain lain. Baru baru ini, peneliti banyak memfokuskan pada task gabungan pada ABSA, yang termotivasi dari manfaat tersedianya referensi yang presisi dan komprehensif bagi pemegang keputusan. Salah satu subtask dari ABSA yaitu ASQE, yaitu Aspect Sentiment Quadruple Extraction. Task ini mengekstrak sentimen quadruple yang terdiri dari aspek, opini, kategori dan polaritas di dalam suatu kalimat.

Task	Review 1: Waiters are very friendly and the pasta is simply average.	Review 2: 效果不错, 第二次购买了, 好评。
ATE	Aspect terms: {Waiters, pasta}	Aspect terms: {效果}
OTE	Opinion terms: {friendly, average}	Opinion terms: {好评, 不错}
AOPE	Pairs: {(Waiters, friendly), (pasta, average)}	Pairs: {(效果, 不错), (__, 好评)}
ASTE	Triples: {(Waiters, friendly, Positive), (pasta, average, Neutral)}	Triples: {(效果, 不错, 正面), (__, 好评, 正面)}
ASQE	Quadruples: {(Waiters, friendly, Service, Positive), (pasta, average, Food, Neutral)}	Quadruples: {(效果, 不错, 功效, 正面), (__, 好评, 整体, 正面)}

**Gambar 2.3** Contoh teks dan ekstraksi quadruple-nya

Pada Gambar 2.3, dua quadruple yang diesktrak (‘waiters’, ‘friendly’, ‘service’, ‘positive’) dan (‘pasta’, ‘average’, ‘food’, ‘neutral’) yang mengklasifikasikan aspek ‘waiters’ dalam kategori ‘service’ dan aspek ‘pasta’ dalam kategori ‘food’. Karena itu, task sentiment annalisis ini lebih lengkap ketimbang task ASTE. (Zhang et al., 2022)

## 2.2.6 Metrik Evaluasi

*F1 score* adalah metrik yang digunakan untuk mengukur klasifikasi berupa nilai diantara 0 dan 1. *F1 -score* adalah rata-rata harmonik dari presisi dan *recall*. *F1 score* sangat bermanfaat dalam situasi di mana kesalahan *false positives* (prediksi salah sebagai positif) dan *false negatives* (tidak mendeteksi positif) memiliki dampak yang seimbang atau penting. Itu membantu untuk memberikan gambaran keseluruhan tentang seberapa baik model klasifikasi mampu memprediksi kelas positif dengan mempertimbangkan keseimbangan antara presisi dan *recall*. (Nugroho, 2019) . Pada persamaan 2.1 ditunjukkan cara perhitungan nilai *F1 score*.

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (2.1)$$

Akurasi adalah parameter pengukuran untuk melihat performa model dalam melakukan klasifikasi. Akurasi merupakan perbandingan dari jumlah prediksi benar dengan total jumlah prediksi model secara keseluruhan. Persamaan 2.2 menunjukkan cara perhitungan nilai akurasi

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative} \quad (2.2)$$

Presisi (*precision*) mengukur seberapa akurat model dalam mengidentifikasi instance positif dari keseluruhan instance yang diprediksi sebagai positif. Metrik ini menggambarkan keakuratan data yang diharapkan dengan hasil prediksi yang dihasilkan model. Persamaan 2.3 menunjukkan rumus perhitungan presisi

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2.3)$$

*Recall* , atau juga dikenal sebagai *sensitivity*, mengukur seberapa baik model dapat mengidentifikasi *instance* positif dari keseluruhan *instance* yang benar-benar positif. Metrik ini menggambarkan tingkat keberhasilan model dalam mengukur seberapa banyak dari prediksi positif yang sebenarnya benar. Persamaan 2.4 menunjukkan perhitungan *recall*

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (2.4)$$

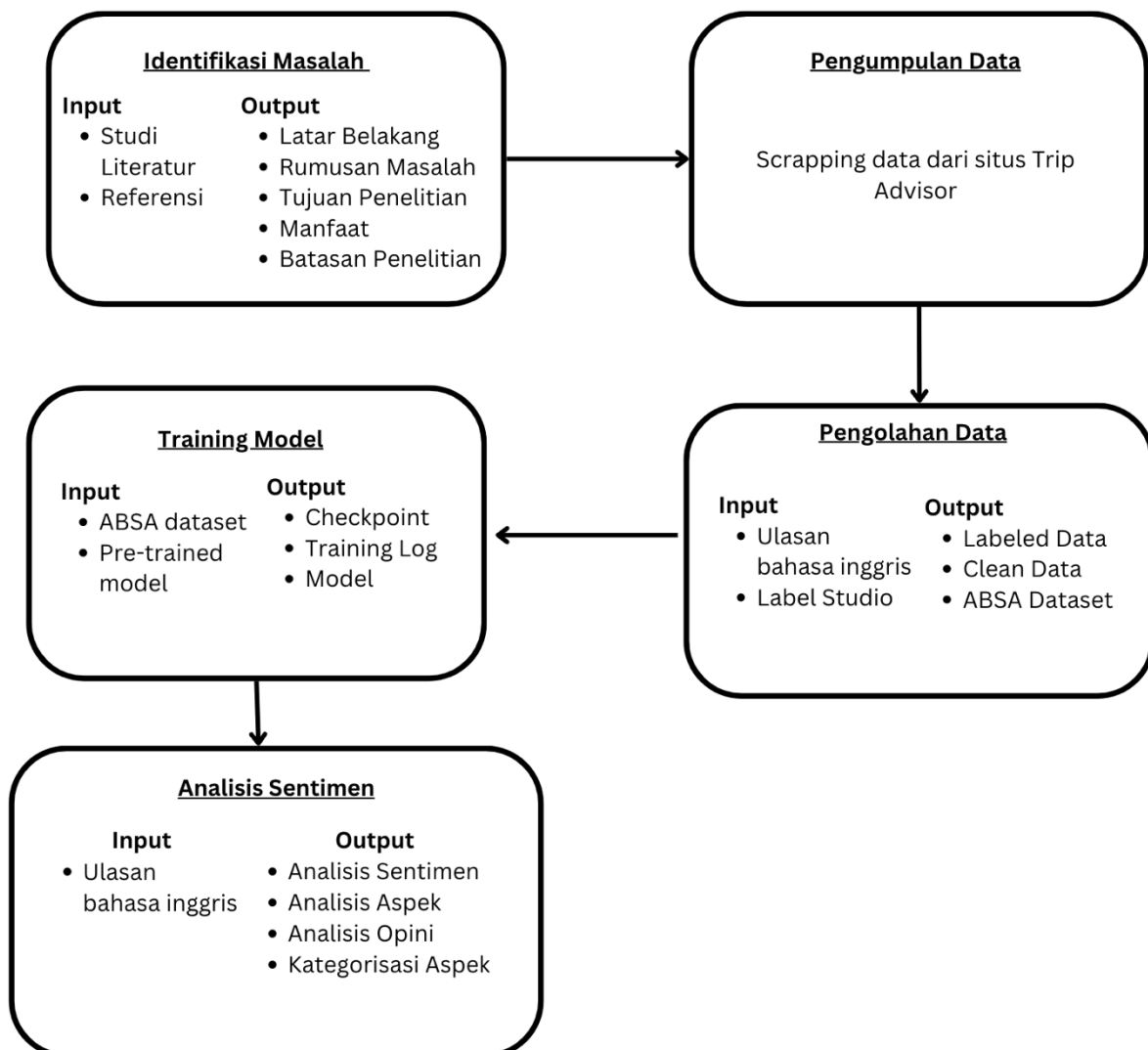


## BAB 3 METODOLOGI

Bab ini menjelaskan tentang metode-metode yang digunakan dalam pengerjaan tugas akhir agar dihasilkan tugas akhir yang sesuai dengan perencanaan dan sesuai target.

### 3.1 Alur Pengerjaan Tugas Akhir

Subbab ini berisi tentang urutan pengerjaan tugas akhir dari awal hingga akhir sehingga pengerjaan dilakukan secara terstruktur.



*Gambar 3.1 Langkah-langkah penelitian*

Gambar 3.1 menunjukkan langkah langkah penelitian yang dimulai dari identifikasi masalah, pengumpulan data dari Trip Advisor, pelatihan model, pengolahan data, serta analisis sentimen.

## 3.2 Urutan pelaksanaan penelitian

Subbab ini membahas tentang langkah-langkah dalam melakukan penelitian

### 3.2.1 Identifikasi Masalah

Penelitian ini dilakukan untuk mengidentifikasi masalah dan penyelesaian masalah dalam sektor pariwisata, khususnya di provinsi Bali. Penelitian ini memerlukan studi literatur untuk menemukan referensi ilmiah. Referensi ilmiah ini didapatkan dari jurnal, buku, web bereputasi, serta penelitian-penelitian terdahulu. Identifikasi metode juga dilakukan untuk menemukan metode dan jenis model yang digunakan. Sumber data untuk penelitian ini diambil dari situs web TripAdvisor dari ulasan 48 hotel dengan total 2108 kalimat yang diambil secara bervariasi untuk memastikan variasi pada polaritas sentimen yang sudah di-*scrapping* dan diolah oleh (Ziaulhaq,2023). Data yang disajikan peneliti tersebut berbentuk triplet dengan 3 label yaitu aspek, sentimen, dan opini. Oleh karena itu, perlu dilakukan pelabelan untuk label kategori untuk memenuhi penelitian ini.

### 3.2.2 Pra-pemrosesan data

Pada tahapan ini dipersiapkan data mentah agar dapat digunakan untuk analisis. Di tahap ini, data teks ulasan akan dibersihkan, diberi label, dan dieksplorasi. Pembersihan data bertujuan untuk menghilangkan kesalahan, ketidakkonsistenan, dan data yang tidak relevan. Ini dilakukan agar data menjadi lebih akurat dan dapat dianalisis lebih efektif. Pelabelan data bertujuan untuk memberikan kategori atau kelas pada data. Hal ini dilakukan agar model dapat mempelajari pola dan hubungan antar data. Eksplorasi data bertujuan untuk memahami data secara mendalam untuk mengidentifikasi pola dan tren yang dapat digunakan untuk meningkatkan akurasi model. Pelabelan menggunakan program Label Studio yang menghasilkan output berupa json. Label dari ulasan ini memiliki empat jenis yaitu aspek, sentimen, opini, dan kategori aspek.

Dengan menggunakan cara yang sama oleh (Harper et al., SemEval 2021) dalam pelabelan, untuk menjaga subjektivitas pelabelan data oleh satu *annotator*, pelabelan dilakukan secara manual oleh dua orang mahasiswa Sistem Informasi Institut Teknologi Sepuluh Nopember Surabaya dan dilanjutkan *review* bersama untuk menyelesaikan perbedaan persepsi dan menjamin konsistensi data. Jika dua orang yang melabeli data menyetujui label data, maka data itu dimasukkan ke dalam dataset. Dalam pelabelan dataset, pelabelan dilakukan sesuai dengan pemahaman manusia yang berarti akan ada faktor subjektivitas. Pelabelan dataset ini menggunakan paradigma deskriptif dalam melabeli *task* NLP yang subjektif (Rottger et al., NAACL 2022), yaitu cara untuk membuat kumpulan data yang mewakili berbagai cara orang menggunakan bahasa. Hal ini dilakukan dengan mendorong anotator untuk mengekspresikan interpretasi dan penilaian mereka sendiri, alih-alih mengikuti aturan yang ketat.

### 3.2.3 Pembuatan Model

Dengan data yang sudah diproses, dibersihkan dan dilabeli, tahapan ini membuat model dengan dataset yang sudah dibagi menjadi *train*, dan *test*. Model akan dilatih berdasarkan dataset ini menggunakan pre-trained model *googleflan-t5-base* dengan inisiasi parameter yang ditemukan pada source code (Yang et al., 2022) seperti learning rate  $5e-5$ , epoch sejumlah 6, dan lain lain. Model dilatih dengan berbagai dataset yang mempunyai beragam kompleksitas seperti opini terdiri dari dua kata berurutan dan terdapat beberapa opini dalam kalimat. Untuk mengetahui berapa hasil yang dapat dianggap cukup, diperlukan baseline model dari berbagai referensi untuk dijadikan acuan.

ACOS dataset (*Aspect-Category-Opinion-Sentiment*) adalah dataset yang digunakan untuk analisis sentimen yang rinci terutama dalam ulasan produk. Ada dua macam ACOS dataset yang dijadikan acuan untuk percobaan dalam model *machine learning* ACOS, yaitu Restaurant-ACOS yang merupakan ekstensi dari SemEval 2016 Restaurant (Xu et al, 2020). Dataset Restaurant-ACOS terdiri dari 3658 quadruples (Cai et al., 2021). Selain itu, dataset ACOS lainnya adalah Laptop-ACOS yang dikumpulkan dari domain Amazon Laptop dengan ukuran dua kali dari Semeval Laptop (Cai et al., 2021). Berikut adalah berbagai hasil model ACOS yang dikumpulkan dari beberapa referensi

**Tabel 3.1 Perbandingan berbagai metode untuk *quadruple task* berdasarkan ACOS-Dataset**

Metode	REST-ACOS			LAPTOP-ACOS			Base
	P	R	F1	P	R	F1	
DP-ACOS Cai et al (2021)	34.67	15.08	21.04	13.0	5.70	8.0	Rule-Based
JET-ACOS Cai et al (2021)	59.81	28.94	39.01	44.52	16.25	23.81	BERT-Based
Extract-Classify-ACOS Cai et al (2021)	38.54	52.96	44.61	45.56	29.48	35.80	
BARTABS(AHoang) Hoang et al (2022)	55.77	50.66	53.09	35.80	38.01	36.88	BART-Based
BARTABS(split) Hoang et al (2022)	56.80	51.09	53.45	41.06	37.89	39.41	
GAS(Bao) Zhang et al (2021c)	60.69	58.52	59.59	41.60	42.75	42.17	T5-Based
SpecialSymbols+UAUL Hu et al (2023)	61.22	59.87	60.53	44.38	43.65	44.01	
SpecialSymbols Hu et al (2022)	59.98	58.40	59.18	43.58	42.72	43.15	
SpecialSymbols+UAUL Hu et al (2023)	61.22	59.87	60.53	44.38	43.65	44.01	
OpinionTree Bao et al (2022)	<b>63.96</b>	<b>61.74</b>	<b>62.83</b>	<b>46.11</b>	<b>44.79</b>	<b>45.44</b>	

Berdasarkan survei pada tabel 3.1, dapat dilihat bahwa model dengan T5-Based cenderung memiliki hasil yang lebih baik ketimbang BART-Based, BERT-Based, maupun Rule-Based. Hasil tertinggi diperoleh dengan F1 sebesar 62.83% pada dataset REST-ACOS dan F1 45.44 pada dataset LAPTOP ACOS.

### 3.2.4 Analisis Sentimen

Setelah pembuatan model sudah final, tahap yang dilakukan selanjutnya adalah membuat *pipeline* yang mengolah teks ulasan dan mengekstraknya menjadi 4 label. Tahap ini berguna

untuk menunjukkan kemampuan model saat disajikan data-data di dunia nyata. Luaran ini dapat berguna untuk pengunjung maupun stakeholder hotel. Dengan mengolah teks ulasan ini menjadi quadruple, data quadruple ini dapat membantu pihak customer dan pengolah hotel dalam menganalisis teks ulasan.

### **3.2.5 Penulisan Tugas Akhir**

Penulisan tugas akhir dilakukan dari awal identifikasi masalah penelitian ini sampai tahap akhir penelitian. Ini dilakukan agar penelitian terdokumentasi dengan baik dan sistematis.

## BAB 4 HASIL DAN PEMBAHASAN

### 4.1 Pra-pemrosesan data

Pada tahap ini, dataset disamakan dalam format dataset referensi utama. Dataset yang akan digunakan berupa jsonl.

```
{
  "text": "Feels like at home . Location is good . Near to sunset road . Friendly staffs . We will stay here for sure when in Bali . Cannot wait and very excited until July 2016 for next bali visit . My kids love to stay here . Minus : 2nd wooden floor is noisy when someone walk will disturb neighbour .",
  "labels": [
    ["Location", "good", "POS", "LOCATION#GENERAL"],
    ["staffs", "Friendly", "POS", "SERVICE#HOSPITALITY"],
    ["2nd wooden floor", "noisy", "NEG", "ROOM#EXPERIENCE"],
    ["My kids", "love to stay", "POS", "HOTEL#EXPERIENCE"]
  ]
}
```

```
{
  "text": "I have stayed there for 3 days with family . Location is very good . Hotel staffs are friendly and provide all helps and recommendation . I would recommend to stay there when you are on solo trip . You can find other better options when you are planning trip to Kuta along with family .",
  "labels": [
    ["Location", "good", "POS", "LOCATION#GENERAL"],
    ["staffs", "friendly", "POS", "SERVICE#HOSPITALITY"],
    ["staffs", "provide all helps", "POS", "SERVICE#HOSPITALITY"],
    ["stay", "recommend", "POS", "HOTEL#GENERAL"]
  ]
}
```

Di format berikut, label label dibungkus sebuah *array* yang setiap labelnya berisi “aspect”, “opinion”, “polarity” dan “category”. Untuk category, terdapat dua bagian, bagian pertama yaitu kategori secara umum, dan bagian kedua adalah konteks yang lebih detail. Untuk menyamakan format dataset, digunakan script dan library *python*.

### 4.2 Skenario dataset

Dalam dataset yang akan digunakan, dataset dibagi menjadi dua skenario, yaitu dataset *multi sentences*, yaitu dataset dengan per 1 teks nya terdapat lebih dari satu kalimat, dan dataset *single sentence*, yaitu dataset dengan per 1 teks nya terdapat satu kalimat saja. Kedua ini berisi dataset yang sama, hanya berbeda jumlah kalimat saja. Berikut merupakan contoh kalimat dalam masing masing dataset.

**Tabel 4.1 Contoh dataset**

Jenis Dataset	Kalimat	Label
Multi Sentence	Beautiful , comfortable and friendly hotel near the beach . Great value for the money . Good situation , next to a big mall and the beach . Rooms really comfortable , clean and no noisy . Wifi excellent .	[           ["money", "Great value", "POS", "HOTEL#PRICE"],           ["hotel", "friendly", "POS", "SERVICE#HOSPITALITY"],           ["hotel", "comfortable", "POS", "HOTEL#EXPERIENCE"],           ["hotel", "Beautiful", "POS", "HOTEL#AMBIENCE"],           ["Rooms", "comfortable", "POS", "ROOM#EXPERIENCE"],           ["Rooms", "clean", "POS", "ROOM#CLEANLINESS"],           ["Rooms", "no noisy", "POS", "ROOM#AMBIENCE"],           ["Wifi", "excellent", "POS", "FACILITY#WIFI"],           ["beach", "next to", "POS", "LOCATION#ATTRACTIONS"],         ]

		["mall", "next to", "POS", "LOCATION#ENTERTAINMENT"]]
Single Sentences	"The bedroom is clean and service is good ."	"labels": [ ["bedroom", "clean", "POS", "ROOM#BED"], ["service", "good", "polarity": "POS", "SERVICE#GENERAL"]]

### 4.3 Pelabelan data

Pada subbab ini dijelaskan tentang langkah-langkah yang dilakukan untuk pelabelan data pada *dataset*.

Data yang digunakan awalnya adalah data yang sudah diambil melalui *web-scraping* dan dilabeli menggunakan *label studio*. Dataset sudah terlabeli tiga komponen triplet (aspek, opini, sentiment). Karena itu, dataset perlu ditambahkan komponen kategori. Kategori yang ditambahkan terdiri dari kategori utama dan sub-kategori. Contoh pada "HOTEL#PRICE", berarti kategori utamanya hotel, dan sub-kategorinya adalah harga. Artinya kategori tersebut menjelaskan tentang hotel pada konteks harga. Pada tabel 4.2 dan 4.3 disebutkan kategori beserta sub-kategori yang dimasukkan ke dalam dataset.

**Tabel 4.2 Daftar Kategori bagian 1**

HOTEL	FOOD	FACILITY	SERVICE
"HOTEL#GENERAL"	"FOOD#GENERAL"	"FACILITY#GENERAL"	"SERVICE#GENERAL"
"HOTEL#PRICE"	"FOOD#TASTE"	"FACILITY#PRICE"	"SERVICE#HOSPITALITY"
"HOTEL#CLEANLINESS"	"FOOD#CHOICE"	"FACILITY#CLEANLINESS"	"SERVICE#AMENITIES"
"HOTEL#EXPERIENCE"	"FOOD#PRICE"	"FACILITY#EXPERIENCE"	"SERVICE#ASSISTANCE"
"HOTEL#AMBIENCE"		"FACILITY#WIFI"	
"HOTEL#SPACE"		"FACILITY#POOL"	
"HOTEL#MAINTAIN"		"FACILITY#MAINTAIN"	

**Tabel 4.3 Daftar Kategori bagian 2**

LOCATION	ROOM
"LOCATION#GENERAL"	"ROOM#GENERAL"
"LOCATION#NEARBY_AMENITIES"	"ROOM#CLEANLINESS"
"LOCATION#ATTRACTIONS"	"ROOM#AMBIENCE"
"LOCATION#ENTERTAINMENT"	"ROOM#EXPERIENCE"
"LOCATION#TRANSPORT_HUB"	"ROOM#BATHROOM"
"LOCATION#CENTRAL_CITY"	"ROOM#BED"
"LOCATION#VIEW"	"ROOM#SPACE"
"LOCATION#MAIN_STREET"	"ROOM#VIEW"
	"ROOM#MAINTAIN"

	"ROOM#AIRCON"
	"ROOM#WINDOW"
	"ROOM#FACILITY"
	"ROOM#AMENITIES"

Kategori ini dibuat setelah menelusuri hal hal yang sering dibahas di dalam teks. Semua label yang dalam dataset menggunakan string mengikuti referensi utama, ini karena model T5 merupakan model *text to text* yaitu model yang dapat menerima input berupa teks. Kategori ini diberikan pada label dengan mempertimbangkan konteks pada teks sesuai dengan subjektivitas *annotator*.

#### 4.4 Eksplorasi data

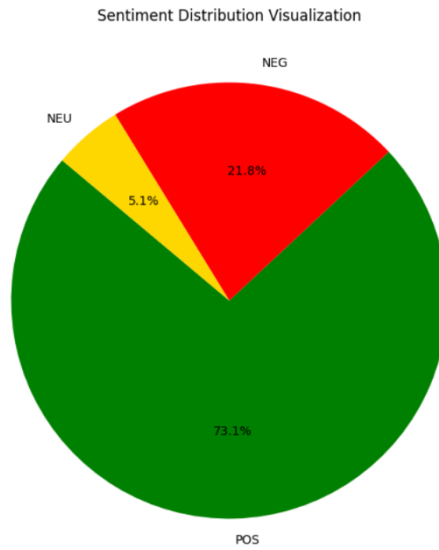
Pada subbab ini dijelaskan eksplorasi data yang dilakukan pada dataset yang digunakan untuk pelatihan model.

Tahap eksplorasi ini dilakukan untuk mengetahui lebih dalam mengenai data serta mendapatkan *insight* dari data. Selain itu, eksplorasi ini juga dapat mengetahui kualitas dari data, seperti apakah data yang ada *balance* atau tidak.

**Tabel 4.4 Jumlah Kalimat dan *quadruple* dataset**

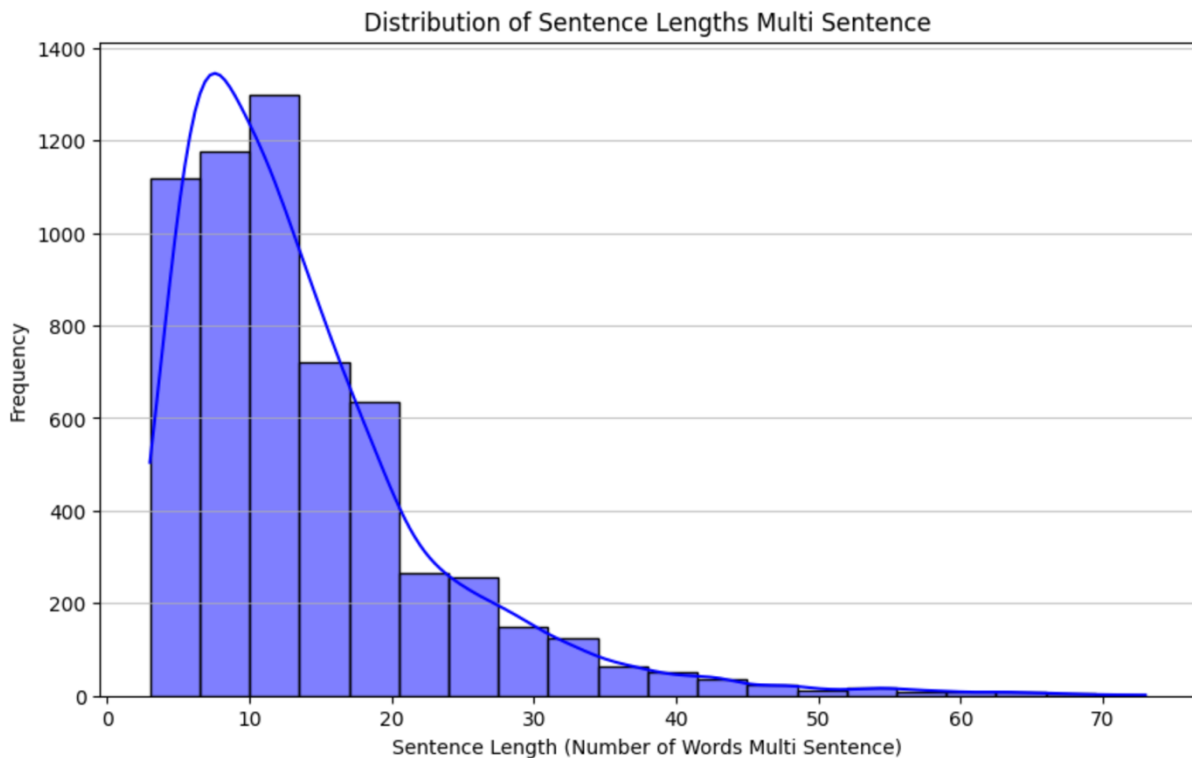
Jumlah Kalimat (Single Sentence)	5977
Jumlah Teks (Multi Sentences)	2102
Jumlah <i>Quadruple</i>	11277

Tabel 4.4 menunjukkan informasi pada dataset yaitu dataset memiliki 5977 kalimat yang ada pada 2102 teks yang masing masing teksnya terdiri dari multi sentences. Di dataset ini terdapat 11277 *quadruple* yang terkestrak dari dataset ini.



**Gambar 4.1 Distribusi Sentimen**

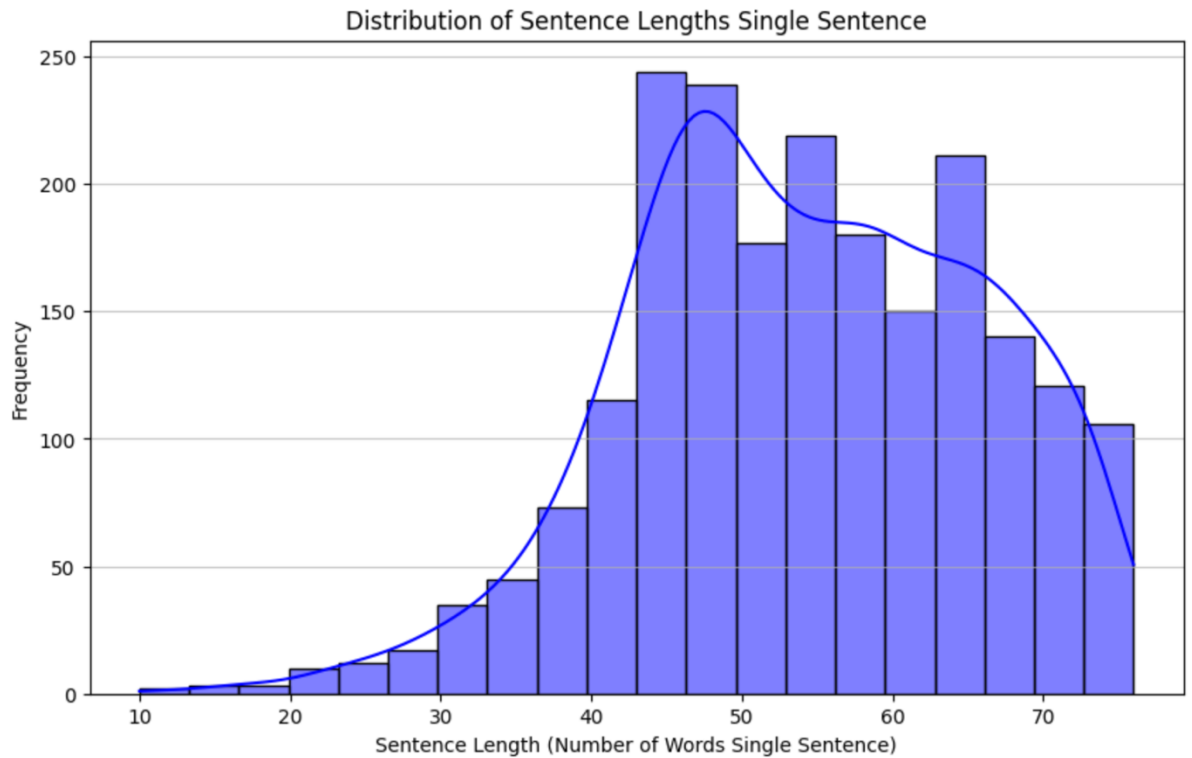
Distribusi sentiment pada dataset pada Gambar 4.1 menunjukkan bahwa mayoritas pada dataset adalah bersentimen positif dengan lebih dari 70% dari total keseluruhan. Salah satu analisis yang dilakukan adalah mengetahui distribusi panjang kalimat pada dataset. Analisis ini bertujuan untuk mengevaluasi distribusi panjang kalimat dan kata pada dataset. Dengan analisis ini, dapat diketahui range, *skewness*, mode, dan *insight* lainnya.



**Gambar 4.2 Distribusi *word length single sentence dataset***

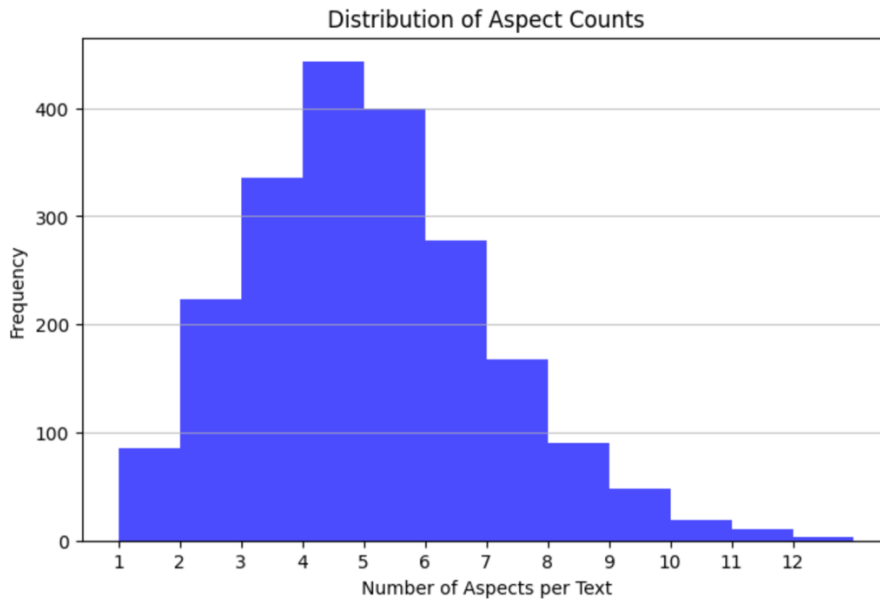


Berdasarkan *Gambar 4.2*, dapat dilihat bahwa distribusi kalimat *right-skewed*, yang berarti kebanyakan kalimat relative pendek, dengan ekor pada grafik yang panjang mengindikasikan kalimat lebih panjang yang lebih sedikit. Panjang kalimat yang paling sering muncul adalah 5-10 kata. Ini menunjukkan kalimat yang ada di dataset single sentence cenderung relatif singkat. Rentang panjang kalimat bervariasi mulai dari 3 hingga 5 kata. Jadi, dapat disimpulkan bahwa dataset ini sebagian besar berisi kalimat yang singkat.

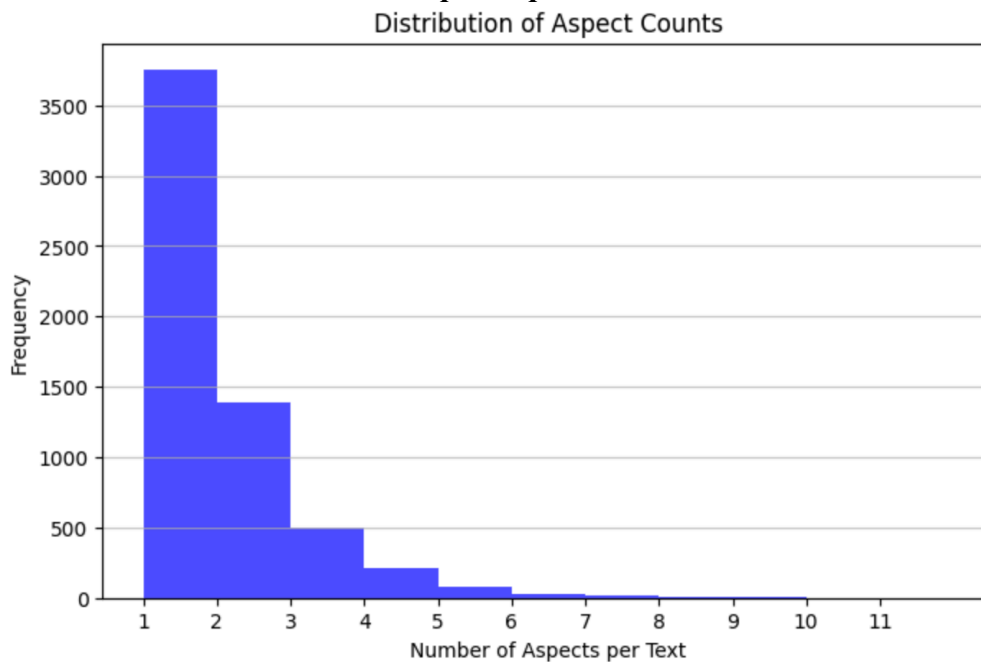


**Gambar 4.3** Distribusi *word length multi sentences dataset*

Berdasarkan *Gambar 4.3*, grafik menunjukkan distribusi normal yang miring ke kanan. Mayoritas kalimat pada dataset memiliki panjang 40 hingga 70 kata, dengan kalimat yang sering muncul memiliki sekitar 40-50 kata yang menjadi puncak dari diagram ini dengan frekuensi mencapai hampir 500. Kalimat di dataset ini memiliki variasi dari rentang 10 hingga 70 kata. Dataset *multi-sentences* ini memiliki perbedaan yang cukup signifikan dengan dataset *single-sentence*, dengan mayoritas kalimat memiliki panjang kalimat yang jauh lebih banyak di dataset *multi sentences* dibanding dataset *single-sentence*.



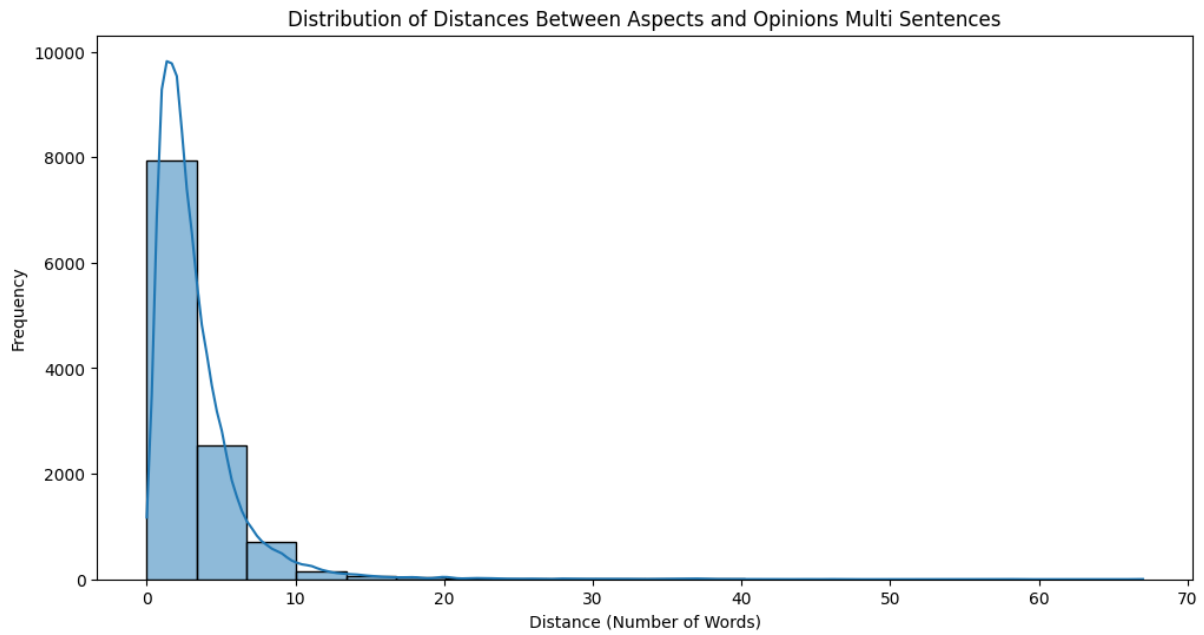
**Gambar 4.4** Distribusi *quadruple multi sentences dataset*



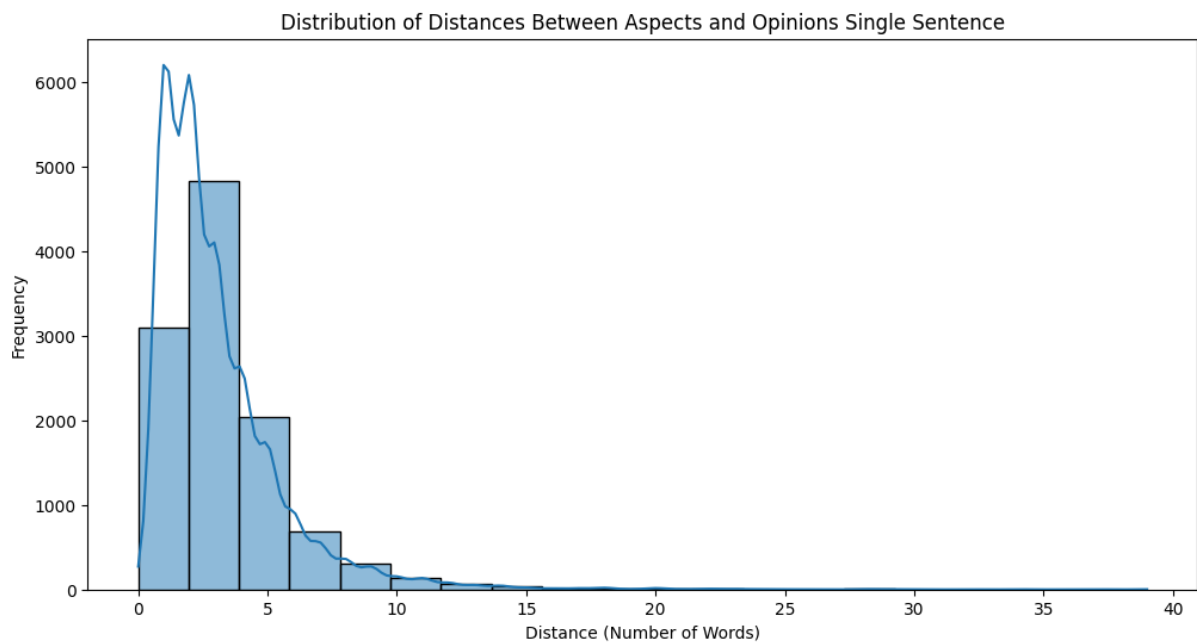
**Gambar 4.5** Distribusi *quadruple single sentence dataset*

Berdasarkan gambar 4.4, *quadruple dataset multi-sentences* berdistribusi normal dengan puncaknya di sekitar 4 hingga 5 *quadruple* per teks dengan distribusinya condong miring ke kanan. Sebagian besar teks memiliki *quadruple* sebanyak 2 hingga 8 *quadruple*. Berbeda dengan dataset *single-sentence* pada gambar 4.5, grafik menunjukkan distribusi normal yang sangat miring ke kanan/ *right skewed* dengan puncaknya di 1 *quadruple*. Ini menunjukkan bahwa teks *multiple sentence* memiliki kompleksitas yang lebih besar ketimbang karena harus mengekstrak *quadruple* yang cenderung lebih banyak.

Selain itu, dilakukan analisis untuk mengetahui jarak antara aspek dan opini, ini perlu diketahui karena salah satu keunggulan arsitektur *Transformers* adalah kemampuan menangkap konteks meskipun jarak antar katanya cukup jauh.

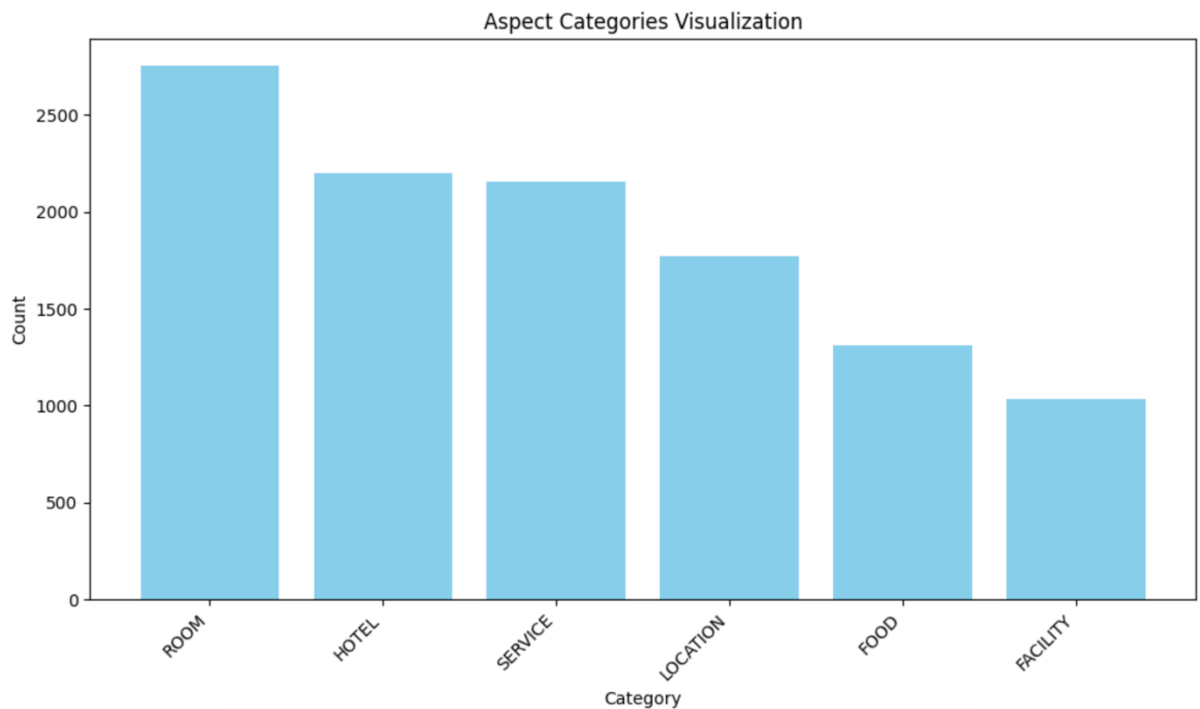


**Gambar 4.6** Distribusi Jarak Aspek dan Opini *multi sentences dataset*



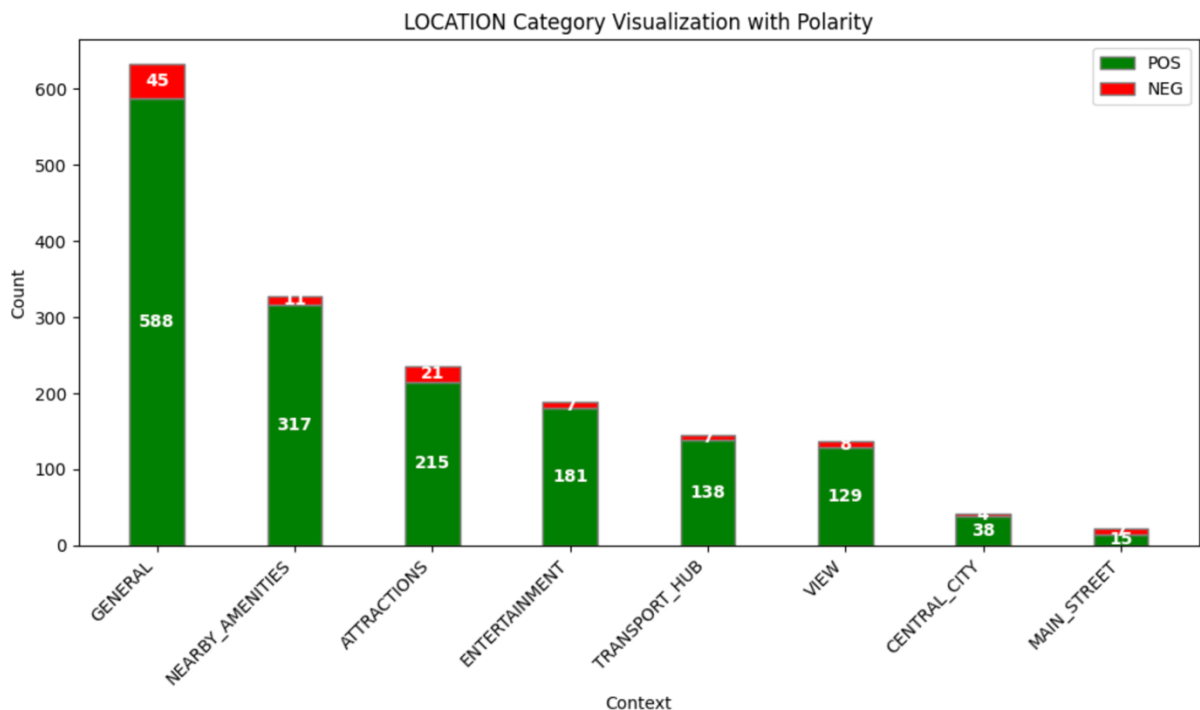
**Gambar 4.7** Distribusi Jarak Aspek dan Opini *multi sentences dataset*

Berdasarkan gambar 4.6 dan 4.7, terlihat bahwa kedua dataset memiliki jarak aspek dan opini yang cenderung dekat, yaitu 0-5 kata. Meskipun begitu, dataset *single-sentence* cenderung memiliki jarak yang lebih dekat ketimbang dataset *multi-sentences*.



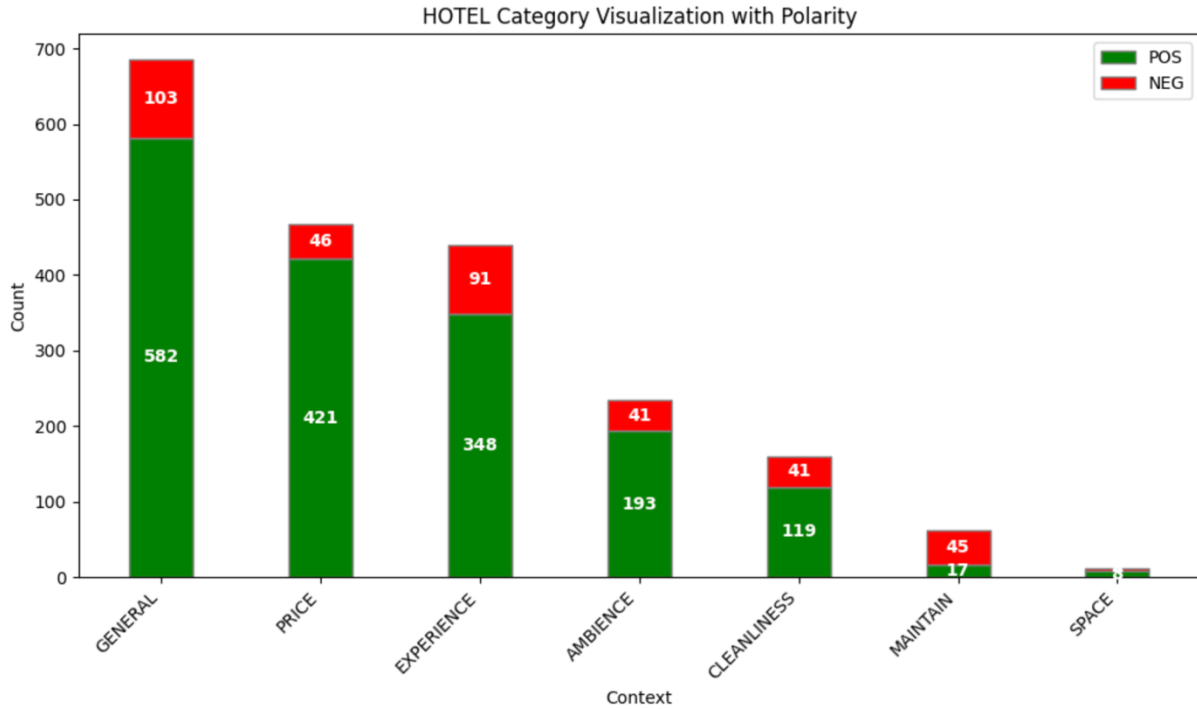
**Gambar 4.8 Visualisasi Distribusi Kategori**

Selanjutnya, dilakukan analisis untuk melihat distribusi kategori, berdasarkan gambar 4.8, distribusi kategori pada dataset relatif cukup seimbang. Kategori yang paling sering dibahas yaitu kategori “ROOM”, dilanjut “SERVICE” dan “HOTEL” yang mempunyai jumlah kategori yang hampir sama. Kategori “FACILITY” dan “FOOD” merupakan dua fasilitas yang paling sedikit dibahas, tetapi secara jumlah cukup sering dibahas.



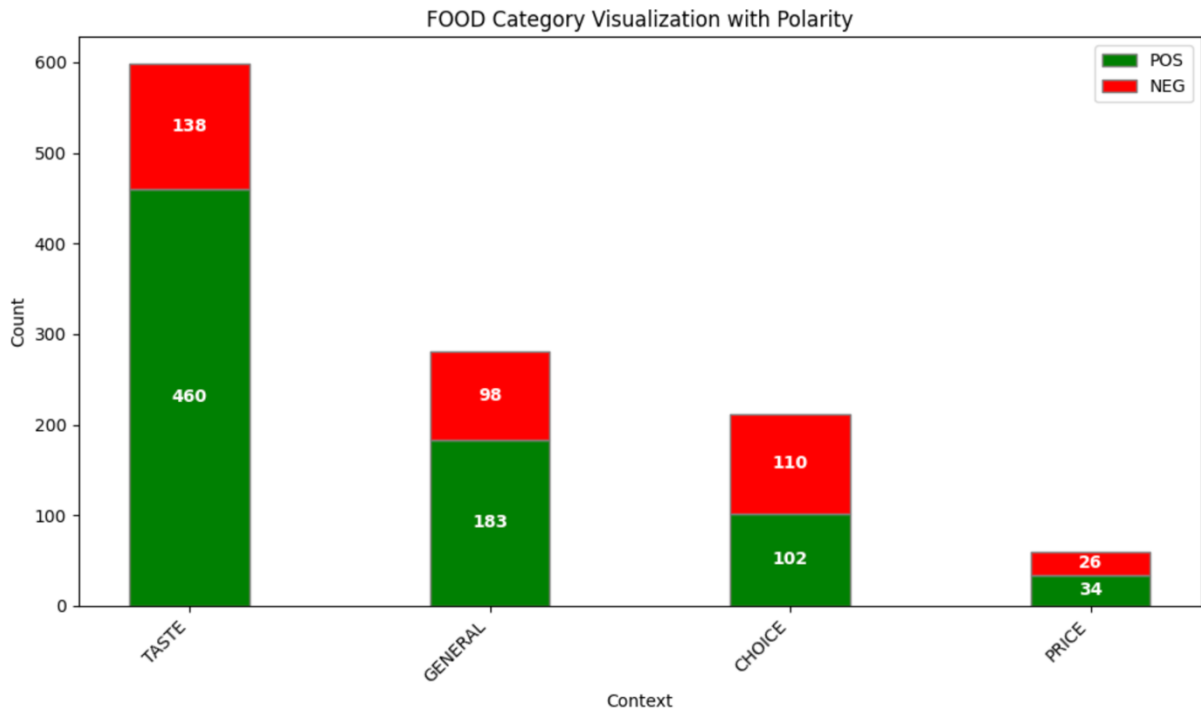
**Gambar 4.9 Visualisasi Distribusi Sub-Kategori Lokasi**

Pada Gambar 4.9, kategori “LOCATION” sub-kategori yang terbanyak yaitu “GENERAL” dengan lebih dari 600 kali disebut dan “NEARBY\_AMENITIES” sebanyak lebih dari 300 kali. Sementara itu, sub kategori “ CENTRAL\_CITY” dan “MAIN\_STREET” paling sedikit dibahas dengan total dari keduanya kurang dari 100. Sebaran polaritas positif dan negatif dalam kategori “LOCATION” terlihat sangat menyimpang dan jauh lebih banyak polaritas positifnya.



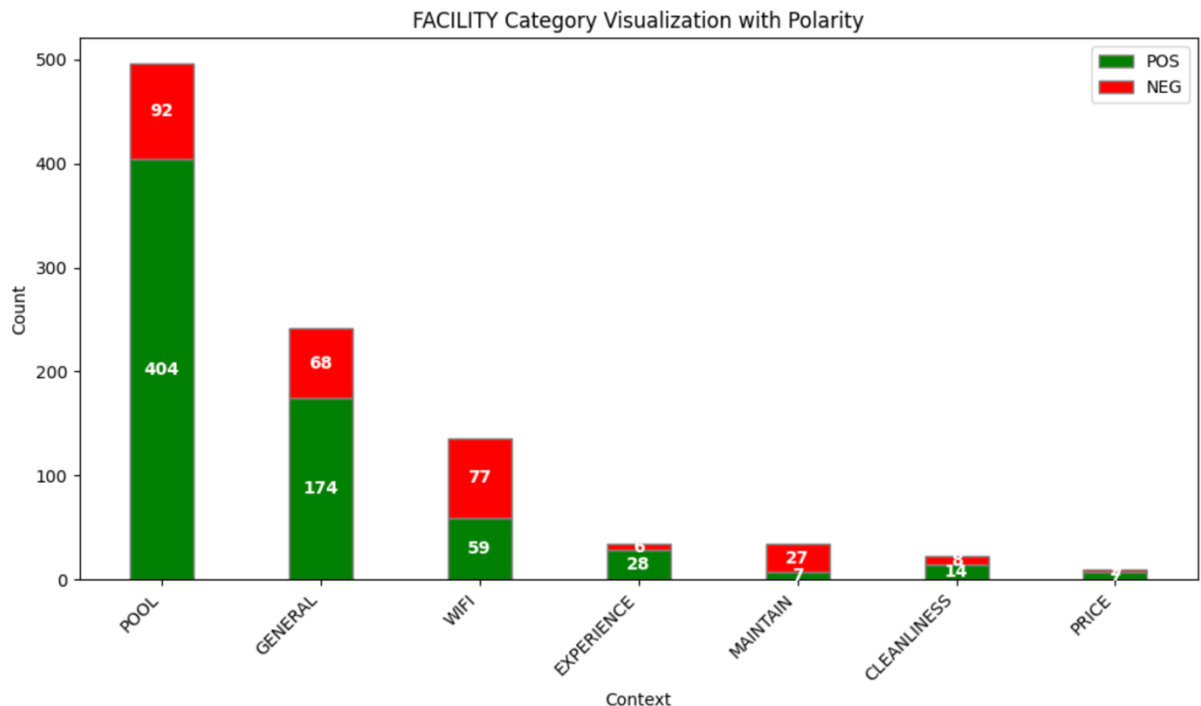
**Gambar 4.10 Visualisasi Distribusi Sub-Kategori Hotel**

Pada gambar 4.10, kategori “HOTEL”, sub kategori terbanyak adalah “GENERAL” dengan hampir 700 kali dilanjut “EXPERIENCE” dan “PRICE” sebanyak lebih dari 450. “MAINTAIN” dan “SPACE” merupakan kategori yang paling sedikit disebut dengan total keduanya kurang dari 100.



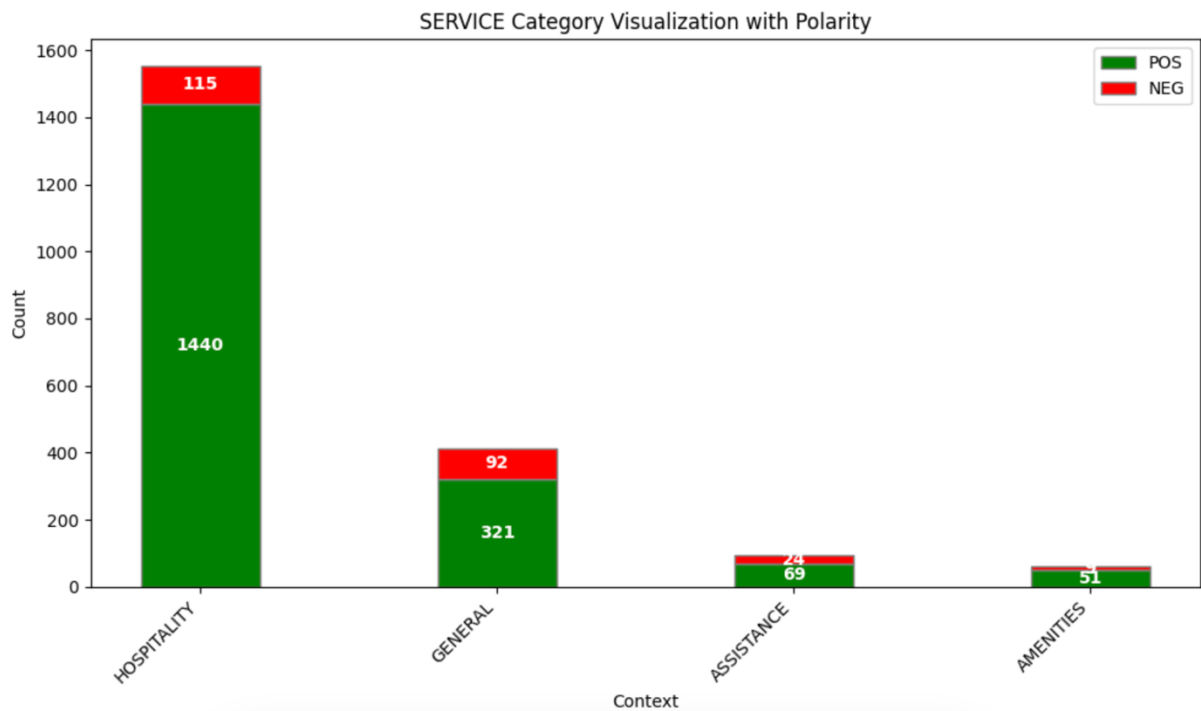
**Gambar 4.11 Visualisasi Distribusi Sub-Kategori "FOOD"**

Berdasarkan Gambar 4.11, data banyak membahas "TASTE" dan "GENERAL" dan "CHOICE". Sub-kategori yang paling banyak dibahas adalah "TASTE" dengan jumlah hampir 600 sentimen.



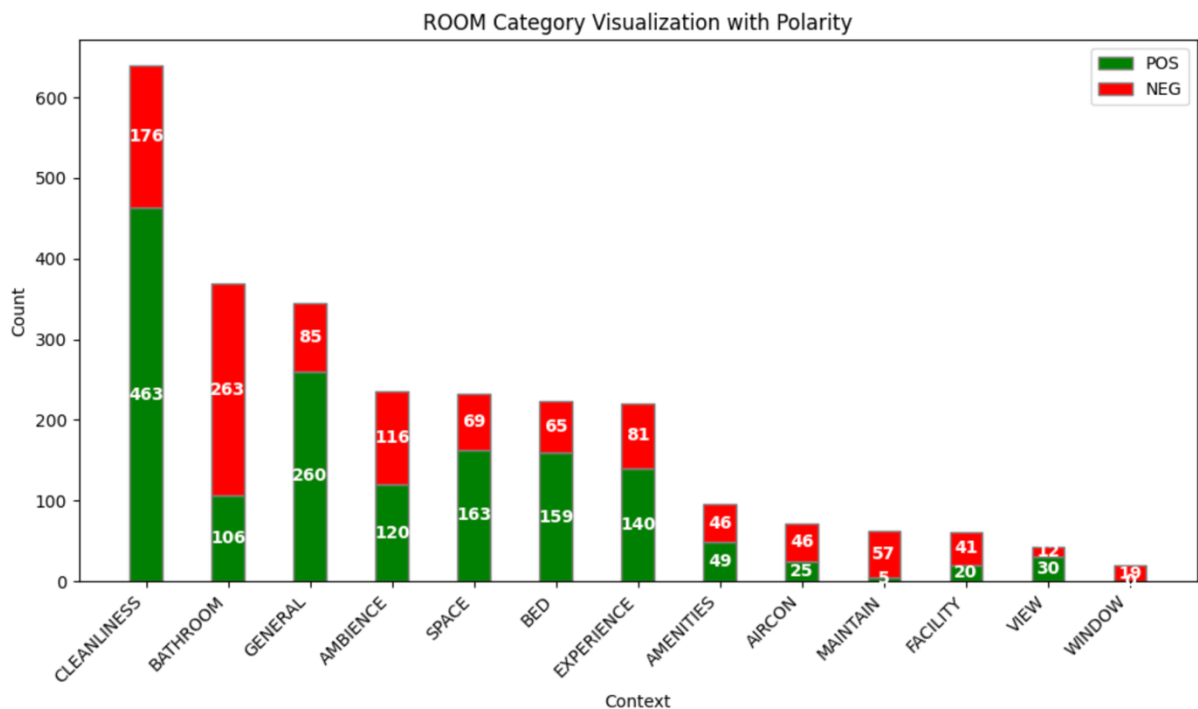
**Gambar 4.12 Visualisasi Distribusi Sub-Kategori "FACILITY"**

Pada Gambar 4.12, kategori "FACILITY", data terlihat cukup tidak seimbang dengan subkategori "EXPERIENCE", "MAINTAIN", "CLEANLINESS", dan "PRICE" menunjukkan proporsi yang jauh lebih sedikit ketimbang "POOL", "GENERAL", dan "WIFI".



**Gambar 4.13 Visualisasi Distribusi Sub-Kategori "FACILITY"**

Pada Gambar 4.13, kategori "SERVICE", data juga terlihat tidak seimbang dengan sebagian besar subkategori terletak pada "HOSPITALITY". Pada kategori ini, antar sub-kategori memiliki kemiripan satu sama lain dan relatif lebih sulit untuk membedakan antar sub-kategori ini.



**Gambar 4.14**

Pada Gambar 4.14, kategori kamar berdasarkan Gambar 4.14 terdapat beberapa sub kategori yang cukup banyak dibahas seperti sub-kategori *cleanliness* yang merupakan sub-kategori yang paling banyak. Beberapa sub-kategori lainnya yang banyak dibahas yaitu *experience*, *cleanliness*, *ambience*, *bed*, *bathroom*, *space*, dan *general*.

## 4.5 Training Data

Pada sub-bab ini dijelaskan mengenai proses training data sampai dengan model hasil *training*

### 4.5.1 Model Baseline

Training data menggunakan model *pre-trained model T5 based* dengan checkpoint yang ada di web hugging-hub. Untuk memperkirakan hasil model yang mau diraih, perlu dilihat bagaimana ng hasil yang diperoleh model dari berbagai metode untuk task *quadruple* ini. Berikut ini adalah hasil dari perbandingan beberapa metode model *benchmark* untuk quadruple task pada ACOS-DATASET

**Tabel 4.5 Baseline Model**

Metode	REST-ACOS			LAPTOP-ACOS			Base
	P	R	F1	P	R	F1	
Caietal (2021)	34.67	15.08	21.04	13.0	5.70	8.0	Rule-Based
JET-ACOS Caietal (2021)	59.81	28.94	39.01	44.52	16.25	23.81	BERT-Based
Extract-Classify-ACOS Caietal (2021)	38.54	52.96	44.61	45.56	29.48	35.80	
BARTABSA(Hoang) Hoangetal (2022)	55.77	50.66	53.09	35.80	38.01	36.88	BART-Based
BARTABSA(split) Hoangetal (2022)	56.80	51.09	53.45	41.06	37.89	39.41	
GAS(Bao) Zhangetal (2021c)	60.69	58.52	59.59	41.60	42.75	42.17	T5-Based
SpecialSymbols +UAUL Huetal (2023)	61.22	59.87	60.53	44.38	43.65	44.01	
SpecialSymbols Huetal (2022)	59.98	58.40	59.18	43.58	42.72	43.15	
SpecialSymbols +UAUL Huetal (2023)	61.22	59.87	60.53	44.38	43.65	44.01	
OpinionTree Baoetal (2022)	<b>63.96</b>	<b>61.74</b>	<b>62.83</b>	<b>46.11</b>	<b>44.79</b>	<b>45.44</b>	

Berdasarkan survey, dapat dilihat bahwa model dengan T5-Based cenderung memiliki hasil yang lebih baik ketimbang BART-Based, BERT-Based, maupun Rule-Based. Hasil tertinggi diperoleh dengan F1 sebesar 62.83% pada dataset REST-ACOS dan F1 45.44 pada dataset LAPTOP ACOS.

### 4.5.2 Persiapan dataset

Data yang sudah diformat dalam bentuk json dibagi menjadi training dan testing dengan perbandingan 70:30. Selanjutnya dataset diubah ke dalam bentuk *dataframe* menggunakan



*library* pandas. Setelah itu, dataframe diconvert menjadi dataset yang sesuai dengan format dataset untuk *training model*. Setelah itu, dataset melalui proses tokenisasi sebelum menjalani *training*. Fungsi-fungsi ini ditulis dalam *python* dalam framework *PyABSA*. Parameter awal untuk pelatihan model ini menggunakan parameter pada referensi (Yang et al., 2021). Berikut merupakan parameter yang digunakan:

```
training_args = {
    "output_dir": model_out_path,
    "evaluation_strategy": "epoch",
    "save_strategy": "epoch",
    "learning_rate": 5e-5,
    "per_device_train_batch_size": 4,
    "per_device_eval_batch_size": 16,
    "num_train_epochs": 3,
    "weight_decay": 0.01,
    "warmup_ratio": 0.1,
    "load_best_model_at_end": True,
    "push_to_hub": False,
    "eval_accumulation_steps": 1,
    "predict_with_generate": True,
    "logging_steps": 1000000000,
    "use_mps_device": False,
    "fp16": False,
}
```

### 4.5.3 Metode Training

Training machine learning model yang menggunakan LLM ini menggunakan *prompt engineering* untuk memberikan instruksi pada model apa inputnya dan apa output yang mesti dikeluarkan. Prompt yang dilakukan adalah memberi instruksi bahwa kalimat inputnya adalah kalimat *review*, dan perintah untuk mengekstrak *quadruple* dari teks. Kemudian model diberikan contoh kalimat inputnya, beserta outputnya. Instruksi selanjutnya adalah memerintahkan model untuk mengekstrak aspek/opini/sentimen/kategori berdasarkan input teksnya. Untuk mengekstrak *quadruple* pada teks, diperlukan 4 fungsi yang masing masing meng-*handle* masing masing task. Fungsi *ATEInstruction*, *APC*, *OpinionInstruction*, dan *CategoryInstruction* masing masing untuk mengekstrak aspek, sentimen, opini, dan kategori. Contoh fungsi dan kelas untuk task dapat dilihat pada lampiran 1. Untuk ekstrak opini, sentimen, dan kategori, pada fungsi juga diinstruksikan aspek mana yang dimaksud lewat parameternya dan outputnya adalah pasangan aspek-opini, aspek-sentimen, dan aspek-kategori. Pada saat training dengan dataset training dan test pada ekstraksi opini, sentimen, dan kategori, parameter aspeknya adalah aspek pada label sebenarnya sehingga ketika aspek yang ditangkap saat prediksi model salah, ekstraksi opini, sentimen, dan kategori ada kemungkinan benar. Ini juga bermanfaat agar dapat diketahui kemampuan model dalam mengekstrak sentimen, opini, dan kategori.

### 4.5.4 Hasil Model

Setelah dilakukan training pada dataset multi sentences dan single sentences, berikut hasil model terbaik yang didapat pada Tabel 4.6:

**Tabel 4.6 Hasil Training**

Dataset	Accuracy	Precision	Recall	F1
<i>Multi-sentences</i>	0.297	0.1745	0.1744	0.1743
<i>Single-sentence</i>	0.702	0.482	0.481	0.479

Hasil prediksi dan label sebenarnya masing-masing disimpan dalam suatu *array*. Dengan urutannya adalah aspek, pasangan aspek-sentimen, pasangan aspek-opini, dan pasangan aspek-kategori. Berikut contoh potongan dari array

**Label prediksi** : ['pool', 'pool:POS', 'pool:amazing', 'pool:FACILITY#POOL', 'staff|cleaning|roomservice', 'staff:POS', 'staff:friendly', 'staff:SERVICE#GENERAL']

**Label sebenarnya** : ['pool', 'pool:POS', 'pool:amazing', 'pool:FACILITY#POOL', 'staff', 'staff:POS', 'staff:friendly', 'staff:SERVICE#GENERAL'].

Hasil model menunjukkan bahwa hasil performa model pada kedua dataset memiliki perbedaan yang cukup jauh. Untuk melihat hasil dari model, perlu diketahui bagaimana performa model untuk memprediksi masing masing task.

**Tabel 4.7 Contoh Prediksi Model MS**

Text	The <b>location</b> is good in the middle of kuta tourist area and room is next to the main road so you will have lot of <b>noise</b> from passing traffic and plenty of <b>cockroaches</b> in the room . The <b>staff</b> is friendly and helpful .
Aspek True	Location, noise, <b>cockroaches</b> , <b>staff</b>
Aspek Prediksi	Location, room, staff
Sentimen True	location:POS, noise:NEG, cockroaches:NEG, staff:POS
Sentimen Prediksi	location:POS, noise:NEG, cockroaches:NEG, staff:POS
Opini True	location:good, noise:lotof, cockroaches:plenty, staff:friendly,staff:helpful
Opini Prediksi	location:good, noise:lot, cockroaches:plenty, staff:friendly, staff:helpful
Kategori True	location:LOCATION#GENERAL noise:ROOM#AMBIENCE cockroaches:ROOM#CLEANLINESS staff:SERVICE#HOSPITALITY staff:SERVICE#HOSPITALITY
Kategori Pred	location:LOCATION#GENERAL noise:ROOM#EXPERIENCE cockroaches:ROOM#EXPERIENCE staff:SERVICE#HOSPITALITY staff:SERVICE#HOSPITALITY
Text	Direct from the Denpasar airport we went to the hotel since we were dead tired after travelling more than 15 hours . Hotel <b>staff</b> were very nice and kind . One employee came and talk to us in our own language . Clean and comfortable enough . Convenient <b>location</b> too. We stay 3 nights in this hotel .
Aspek True	Staff, Hotel, <b>location</b>
Aspek Prediksi	Staff, location
Sentimen True	staff:POS, Hotel:POS, location:POS

Sentimen Prediksi	staff:POS, Hotel:POS, location:POS
Opini True	staff:nice, staff:kind, Hotel:Clean, Hotel:comfortable, location:Convenient
Opini Prediksi	staff:nice, staff:kind, Hotel:Clean, Hotel:comfortable, location:Convenient
Kategori True	staff:SERVICE#HOSPITALITY staff:SERVICE#HOSPITALITY Hotel:HOTEL#CLEANLINESS Hotel:HOTEL#EXPERIENCE location:LOCATION#GENERAL
Kategori Pred	staff:SERVICE#HOSPITALITY staff:SERVICE#HOSPITALITY Hotel:HOTEL#CLEANLINESS Hotel:HOTEL#EXPERIENCE location:LOCATION#GENERAL
Teks	This <b>hotel</b> is a great stay in Legian , Bali , offering great <b>service</b> and clean <b>rooms</b> and is in a very convenient <b>location</b> . Very convenient for shopping , as well as nearby <b>Spas</b> and <b>beach clubs</b> . The reception <b>staff</b> cannot do enough for you with attention to every detail .
Aspek True	hotel, service, rooms, location, shopping, Spas, beachclubs, staff
Aspek Prediksi	hotel, service, rooms, location, shopping, Spas, beachclubs, <b>receptionstaff</b>
Sentimen True	hotel:POS, service:POS, rooms:POS, location:POS, shopping:POS, Spas:POS, beachclubs:POS, staff:NEG
Sentimen Prediksi	hotel:POS, service:POS, rooms:POS, location:POS, shopping:POS, Spas:POS, beachclubs:POS, staff: <b>POS</b>
Opini True	hotel:great, service:great, rooms:clean, location:convenient, shopping:convenient, Spas:nearby, beachclubs:nearby, staff:cannotdoenough
Opini Prediksi	hotel:great, service:great, rooms:clean, location:convenient, shopping:convenient, <b>Spas:convenient,beachclubs:convenient, staff:candoenough</b>
Kategori True	hotel:HOTEL#GENERAL service:SERVICE#GENERAL rooms:ROOM#CLEANLINESS location:LOCATION#GENERAL shopping:LOCATION#NEARBY_AMENITIES Spas:LOCATION#NEARBY_AMENITIES beachclubs:LOCATION#NEARBY_AMENITIES staff:SERVICE#ASSISTANCE
Kategori Pred	hotel:HOTEL#GENERAL service:SERVICE#GENERAL rooms:ROOM#CLEANLINESS location:LOCATION#GENERAL shopping:LOCATION#NEARBY_AMENITIES Spas:LOCATION#NEARBY_AMENITIES

	beachclubs:LOCATION#NEARBY_AMENITIES staff:SERVICE#HOSPITALITY
--	---

Dapat dilihat dari hasil prediksi model pada Tabel 4.7, sebenarnya pada kedua teks model dapat mengekstrak *quadruple* dengan sangat baik namun ada beberapa kesalahan yang dilakukan model. Pada teks ketiga, prediksi model adalah “hotel, service, rooms, location, shopping, Spas, beachclubs, receptionstaff ”. Model memprediksi “reception staff” alih-alih “staff” yang merupakan label asli. Ini terhitung salah karena katanya berbeda tetapi kedua kata ini memiliki arti yang sama. Model menangkap aspeknya sebanyak dua kata, yaitu “reception” dan “staff”. Pada prediksi sentimen, model memprediksi sentimen “staff” sebagai negatif sementara itu label sebenarnya adalah positif. Tetapi setelah melihat teks, kalimatnya adalah “reception staff cannot do enough” yang artinya staff membantu lebih dari cukup. Jadi sentimen yang sebenarnya adalah positif (label sebenarnya salah). Meskipun hasil metrik dari model tidak cukup bagus, sebenarnya model dapat meng-ekstrak *quadruple* dengan cukup bagus. Akan tetapi, karena aspek yang ditangkap teks pada dataset ini banyak, maka jika salah satu saja prediksi nya salah maka akan terhitung salah.

Pada dataset single sentence dilakukan training dengan parameter yang sama, hanya berbeda dataset saja. Berdasarkan analisis, dataset *single sentence* memiliki aspek yang cenderung jauh lebih sedikit ketimbang dataset *multi sentences*. Setelah dilakukan training, terdapat peningkatan cukup signifikan dalam hasil performa model.

**Tabel 4.7 Contoh Prediksi Model SS**

Text	We moved to this <b>Hotel</b> because its closer to the Double Six <b>Beach</b> .
Aspek True	Hotel Beach
Aspek Prediksi	DoubleSixBeach
Sentimen True	Hotel:NEU Beach:POS
Sentimen Prediksi	Hotel:POS Beach:POS
Opini True	Hotel:moved, Beach:closer
Opini Prediksi	Hotel:closer to the Double Six Beach, Beach:closer
Kategori True	Hotel:LOCATION#GENERAL Beach:LOCATION#VIEW
Kategori Pred	Hotel:HOTEL#GENERAL Beach:LOCATION#ATTRACTIONS
Teks	The <b>hotel</b> is brand new so its clean , bright but come with no <b>breakfast</b> but is not big problem since many <b>restaurant</b> around the hotel .
Aspek True	hotel breakfast restaurant
Aspek Prediksi	hotel breakfast
Sentimen True	hotel:POS breakfast:NEG restaurant:POS
Sentimen Prediksi	hotel:POS breakfast:NEG restaurant:POS
Opini True	hotel:brandnew, hotel:clean, hotel:clean, breakfast:no, restaurant:many
Opini Prediksi	hotel:new, hotel:clean, breakfast:no, restaurant:noproblem
Kategori True	hotel:HOTEL#GENERAL, hotel:HOTEL#CLEANLINESS, hotel:HOTEL#CLEANLINESS, breakfast:FOOD#GENERAL, restaurant:FOOD#CHOICE
Kategori Pred	hotel:HOTEL#AMBIENCE,

<p>hotel:HOTEL#CLEANLINESS,  breakfast:FOOD#GENERAL,  restaurant:LOCATION#NEARBY_AMENITIES</p>
--

Pada teks pertama pada Tabel 4.8, label aspek sebenarnya adalah “Hotel” dan “Beach”, sedangkan model memprediksi ,”Hotel” dan “Double Six Beach”. Meskipun berbeda kata, sebenarnya model menangkap aspek yang tepat, yaitu “Double Six Beach”. Untuk opini prediksi model juga tidak salah meskipun berbeda dengan label sebenarnya. Pada teks ketiga, model tidak menangkap aspek “restaurant” yang seharusnya ditangkap. Pada opini, model juga sebenarnya tidak salah seperti memprediksi opini hotel “new” alih alih “brand new” seperti label sebenarnya.

#### 4.5.5 Optimasi

Setelah menganalisis hasil prediksi model, terdapat beberapa kategori yang mirip dan sering misprediksi oleh model. Berikut adalah mapping untuk mengganti beberapa kategori untuk menyederhanakan kategori.

```
category_mapping = {
    "SERVICE#HOSPITALITY" : "SERVICE#GENERAL",
    "HOTEL#EXPERIENCE" : "EXPERIENCE,AMBIENCE",
    "HOTEL#AMBIENCE" : "EXPERIENCE,AMBIENCE",
    "FACILITY#EXPERIENCE" : "EXPERIENCE,AMBIENCE",
    "ROOM#AMBIENCE" : "EXPERIENCE,AMBIENCE",
    "ROOM#EXPERIENCE" : "EXPERIENCE,AMBIENCE",
    "SERVICE#AMENITIES" : "SERVICE#GENERAL",
    "SERVICE#ASSISTANCE" : "SERVICE#GENERAL",
}
```

Setelah menyederhanakan kategori lalu melakukan train *model*, berikut adalah hasil terbaik dari training model pada Tabel 4.8 :

**Tabel 4.8 Hasil Model SS setelah optimasi kategori**

Dataset	Accuracy	Precision	Recall	F1
<i>Hotel-ss</i>	0.769	0.558	0.559	0.557

**Tabel 4.9 Performa model SS per *task***

Dataset	Accuracy	Precision	Recall	F1
<i>Overall</i>	0.769	0.558	0.559	0.557
<i>Aspect</i>	0.7307	0.4625	0.4537	0.4561
<i>Sentiment</i>	0.8628	0.6985	0.7016	0.6986
<i>Opinion</i>	0.7157	0.5501	0.5503	0.5487
<i>Category</i>	0.6368	0.6384	0.6368	0.6331

Tabel 4.9 menyatakan performa model dalam setiap *task*-nya, yaitu aspek, sentimen, opini, dan kategori.

#### 4.5.6 Pengkategorian

Perlu diingat bahwa penggunaan kategori ini dilakukan agar informasi yang di-ekstrak pada data dengan bantuan *model machine learning* dapat dicerna lebih mudah ketimbang hanya

menggunakan aspeknya saja. Jika aspek yang ada pada data terlalu banyak, akan sulit untuk mengakses informasi pada data karena banyaknya aspek yang tersebar. Dengan adanya kategori ini, aspek-aspek akan dikelompokkan menjadi kelompok yang jauh lebih kecil sehingga memudahkan untuk mengambil *insight* dari data. Pada saat melakukan pelabelan kategori, kategori didasarkan pada konteks pada kalimat, tidak hanya berdasarkan aspek saja. Jadi kategori yang dimaksud bukan pengkategorian aspek saja, tetapi mempertimbangkan opini dan konteks pada kalimat secara menyeluruh. Karena penelitian ini cukup menekankan pada pengkategorian, maka perlu diukur performa model di bidang ini. Pada bidang pengkategorian, model memiliki performa sebagai berikut sesuai pada Tabel 4.10:

**Tabel 4.10 Performa model pada kategori utama**

Accuracy	Precision	Recall	F1
0.6368	0.6384	0.6368	0.6331

Karena kategori ini terdiri dari 2 bagian, yaitu kategori utama dan subkategori, perlu diketahui juga performa model pada kategori utama tanpa subkategori pada Tabel 4.11 :

**Tabel 4.11 Performa model pada sub-kategori**

Accuracy	Precision	Recall	F1
0.7488	0.7720	0.7488	0.7553

Pada sub-bab selanjutnya akan dilakukan implementasi penggunaan kategorisasi untuk mengambil *insight* dari dataset.

#### 4.5.7 Perbandingan antar model

Meskipun hasil dari model yang ditrain dari dataset *multi-sentences* secara metrik tidak sebagus model yang dilatih model *single-sentence*, perlu dilakukan analisis kedua model. Selain itu, perlu dilakukan analisis apakah model yang dilatih menggunakan dataset *single sentence* dapat melakukan prediksi pada kalimat yang panjang seperti kalimat yang ada pada dataset *multi sentences*. Karena itu, contoh data yang diuji adalah data dengan teks yang panjang.

**Tabel 4.12 Perbandingan contoh prediksi model ss dan model ms**

Teks	Like the name suggests, the <b>place</b> as indeed very artsy. there was a huge gallery that also served as a restaurant that had just opened, very beautiful. \r\nOur <b>room</b> itself was so aesthetically pleasing, ,so warm & vibrant. our <b>pool</b> was on the edge, overlooking the forest. absolutely beautiful. \r\n <b>Staffs</b> were extremely engaging & helpful.
Prediksi model ms	['place','POS','artsy','HOTEL#AMBIENCE'], ['room','POS','aesthetically pleasing','ROOM#AMBIENCE'], ['pool','POS','warm','ROOM#AMBIENCE'], ['Staffs','POS','vibrant','FACILITY#POOL']
Prediksi model ss	['place','POS','artsy','EXPERIENCE,AMBIENCE'], ['room','POS','esthetically pleasing','ROOM#GENERAL'], ['pool','POS','on the edge','FACILITY#POOL'], ['Staffs','POS','engaged','SERVICE#GENERAL']

Pada tabel, kedua model menangkap empat aspek yang sama yaitu ‘aspek’, ‘room’, ‘pool’, dan ‘Staffs’. Ada beberapa aspek-opini yang tidak ditangkap kedua model, seperti *gallery*,

*restaurant, warm, vibrant, forest*. Pada model *ms*, terdapat opini yang bukan merupakan pasangan dari aspeknya. Ini merupakan salah satu contoh dari banyak contoh di mana model *ms* salah memprediksi pasangan aspek-opininya dan tidak memahami lebih dalam mengenai arti dan konteks sebenarnya dalam kalimat. Model *ss* mempunyai kapabilitas yang lebih mendalam di bidang ini. Karena itu, model *ss* mempunyai performa yang lebih baik ketimbang model *ms* dalam memahami teks dan memprediksi *quadruple*.

## 4.6 Implementasi Model dan Penggunaan Kategorisasi

Pada subbab ini didokumentasikan implementasi dari model untuk mengambil *insight* dari data.

### 4.6.1 Dataset

Data yang digunakan pada implementasi ini didapat dari *scrapping* pada web Booking.com dengan salah satu hotel di Bali. Rating hotel ini di Booking.com adalah 7.7/10. Dibandingkan hotel-hotel di Bali lainnya di situs Booking.com, hotel ini mempunyai rating yang relatif rendah. Salah satu tujuan menggunakan review dari hotel ini yaitu untuk menguji dan mengimplementasikan model *quadruple* untuk meng-*handle* ulasan ulasan negatif mengingat dataset training mempunyai sedikit ulasan negatif. Jumlah ulasannya sebanyak 581 ulasan.

### 4.6.2 Pre-processing

Hasil output *scrapping* berupa text csv yang kemudian di-*convert* ke dalam bentuk dataframe menggunakan *library pandas*. Kemudian semua data diiterasi dan diprediksi *quadruple* nya lalu disimpan ke dalam csv. Untuk dilakukan analisis, dataframe yang berisi teks dan *quadruple* akan divisualisasikan.

### 4.6.3 Analisis data

Setelah diubah ke dataframe, dilakukan analisis data menggunakan *library* di python untuk memvisualisasikan data dengan memanfaatkan hasil *quadruple* teks ulasan yang telah diprediksi oleh model machine learning *quadruple* terbaik yang telah di-*train* peneliti.

#### 4.6.3.1 Contoh prediksi data

Berikut adalah contoh hasil prediksi data oleh model pada Tabel 4.13:

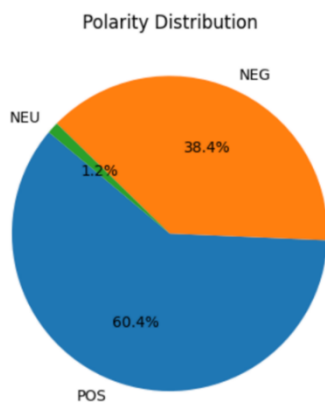
**Tabel 4.13 Contoh implementasi prediksi *model machine learning***

Text	Loved the <b>view</b> , the peace and quiet, and the helpfulness of the <b>staff</b> . <b>Room</b> cleaned every day.
Kategori Pred	[{'view', 'POS', 'Loved', 'LOCATION#VIEW'}, {'staff', 'POS', 'helpfulness', 'SERVICE#GENERAL'}, {'Room', 'POS', 'cleaned every day', 'ROOM#CLEANLINESS'}]
Text	The <b>WiFi</b> was hit or miss, there were times that we couldn't connect or the <b>speeds</b> were slow, and in the breakfast area you didn't catch a signal . The <b>pool</b> could use some reconditioning, there's some moss and some of the kit came out, don't get me wrong, we stayed in the pool more than 6 hours in our last day, it's really great.
Label	[{'WiFi', 'NEG', 'hit or miss', 'FACILITY#WIFI'}, {'speeds', 'NEG', 'slow', 'FACILITY#WIFI'}, {'pool', 'POS', 'could use some reconditioning', 'FACILITY#POOL'}]1

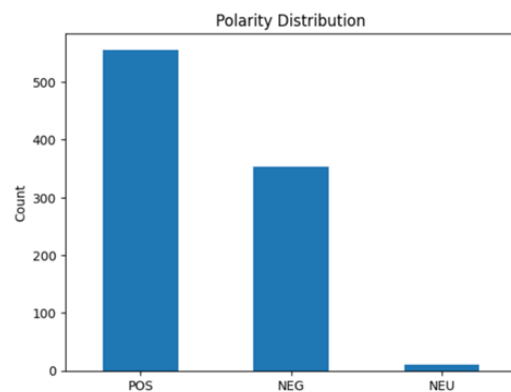
Teks	"The resort is in a tranquil and beautiful <b>location</b> . The 2 <b>pools</b> both have a great <b>view</b> and <b>sitting</b> areas were very relaxing. <b>Breakfast</b> , lunch and <b>dinners</b> were all good with fresh food and nice serving. <b>Staff</b> was wonderful and bent over backward. We missed the tour of the galleries but heard it was amazing!","
Quadruple	['location', 'POS', 'beautiful', 'LOCATION#GENERAL'], ['pools', 'POS', 'great view', 'FACILITY#POOL'], ['view', 'POS', 'sitting areas', 'LOCATION#VIEW'], ['sitting areas', 'POS', 'good', 'EXPERIENCE,AMBIENCE'], ['Breakfast', 'POS', 'good', 'FOOD#TASTE'], ['dinners', 'POS', 'wonderful', 'FOOD#TASTE'], ['Staff', 'POS', 'ank over backward', 'SERVICE#GENERAL']

#### 4.6.3.2 Sebaran polaritas

Sub kategori ini menjelaskan tentang perbandingan sentimen pada dataset



**Gambar 4.15 Pie Chart**  
Distribusi Polaritas



**Gambar 4.16 Bar Chart** Distribusi Polaritas

**Tabel 4.14 Jumlah Sentimen Positif, Negatif, dan Netral**

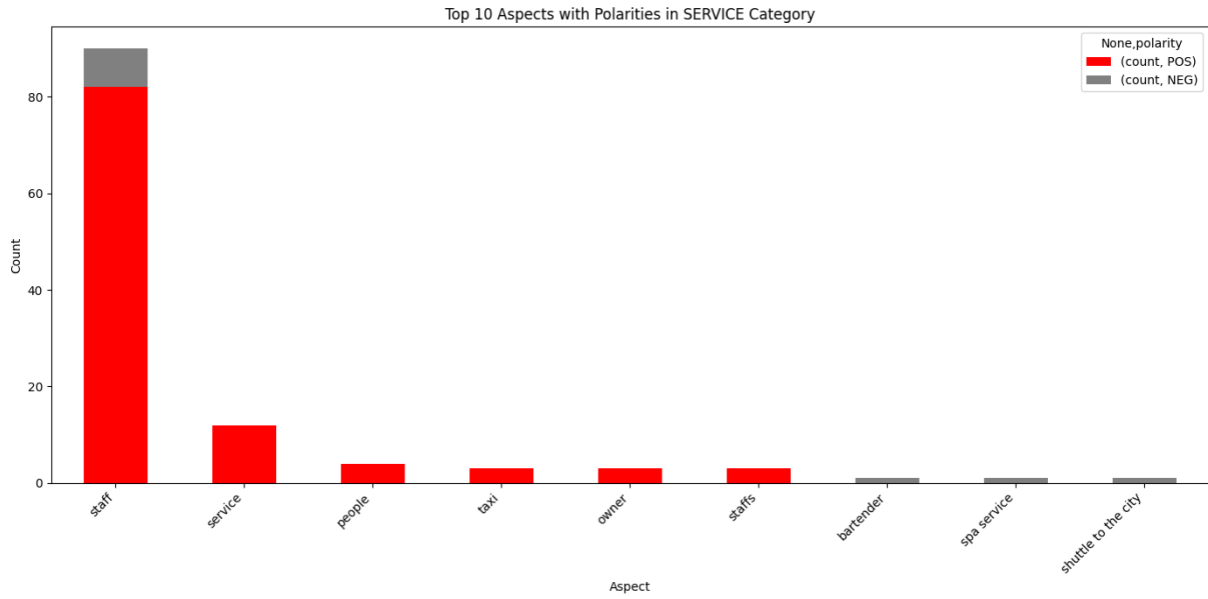
Positif	Negatif	Netral
556	335	11

Berdasarkan diagram pada 4.15, 4.16 dan Tabel 4.14, Mayoritas data bersentimen positif dengan lebih 60.4% nya positif. Akan tetapi, sentimen negatif pada review hotel tanah-merah memiliki proporsi yang cukup besar, yaitu 38,4%. Ini menarik untuk dianalisis lebih mendalam.

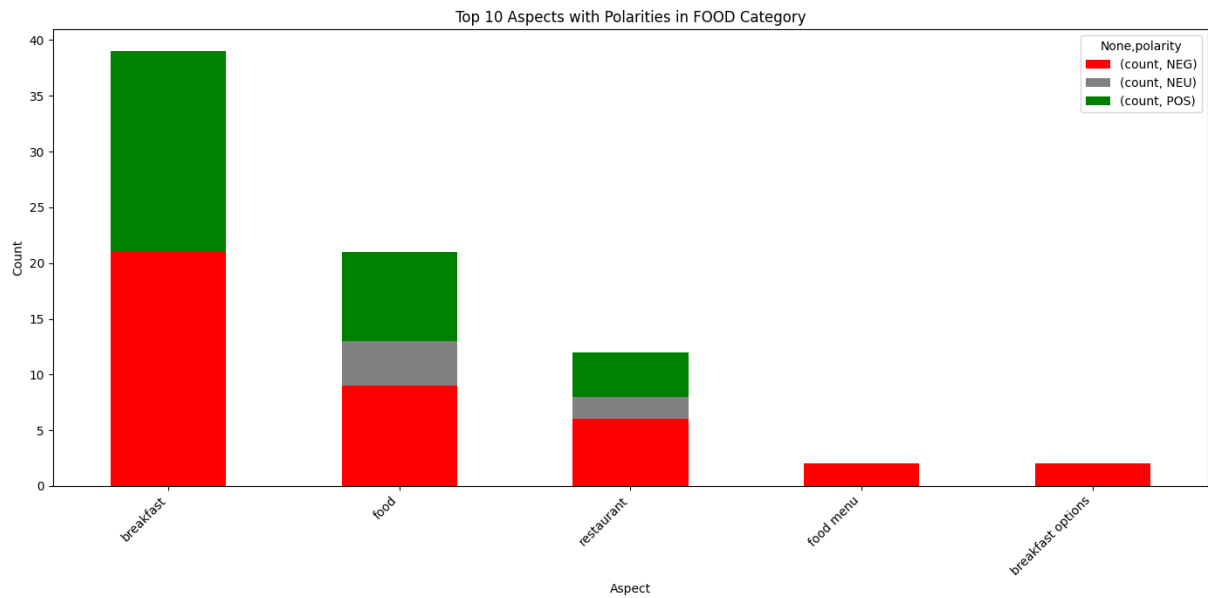
#### 4.6.3.3 Sebaran aspek

Untuk mengetahui apa saja aspek-aspek yang dibahas pada masing masing kategori, peneliti melakukan visualisasi urutan aspek yang sering dibahas per-kategori nya.

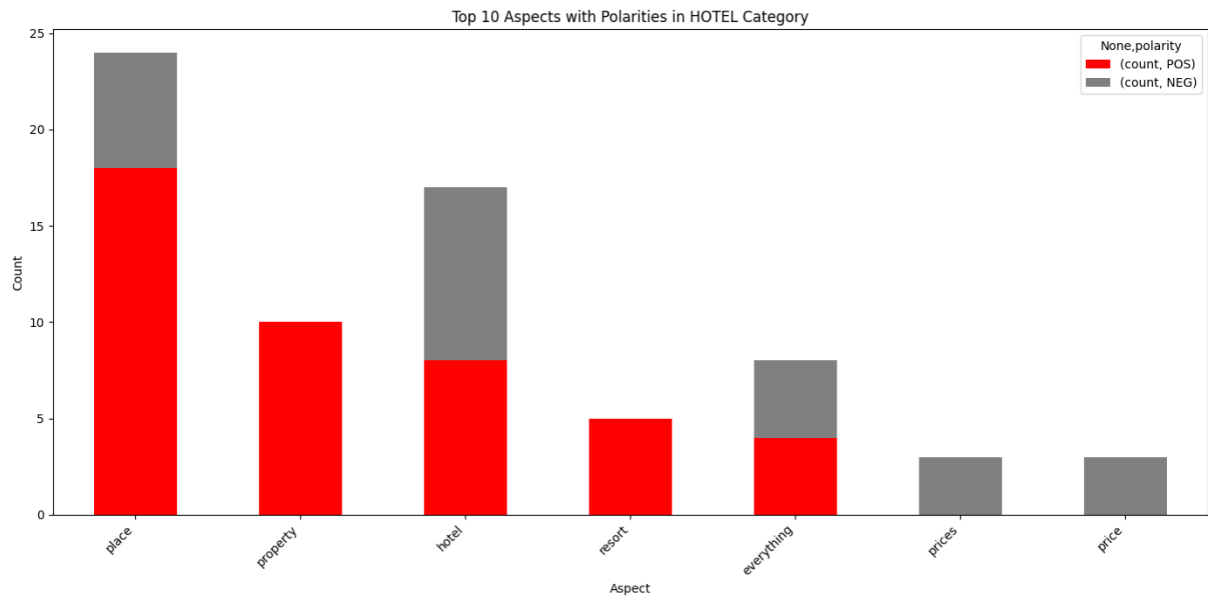




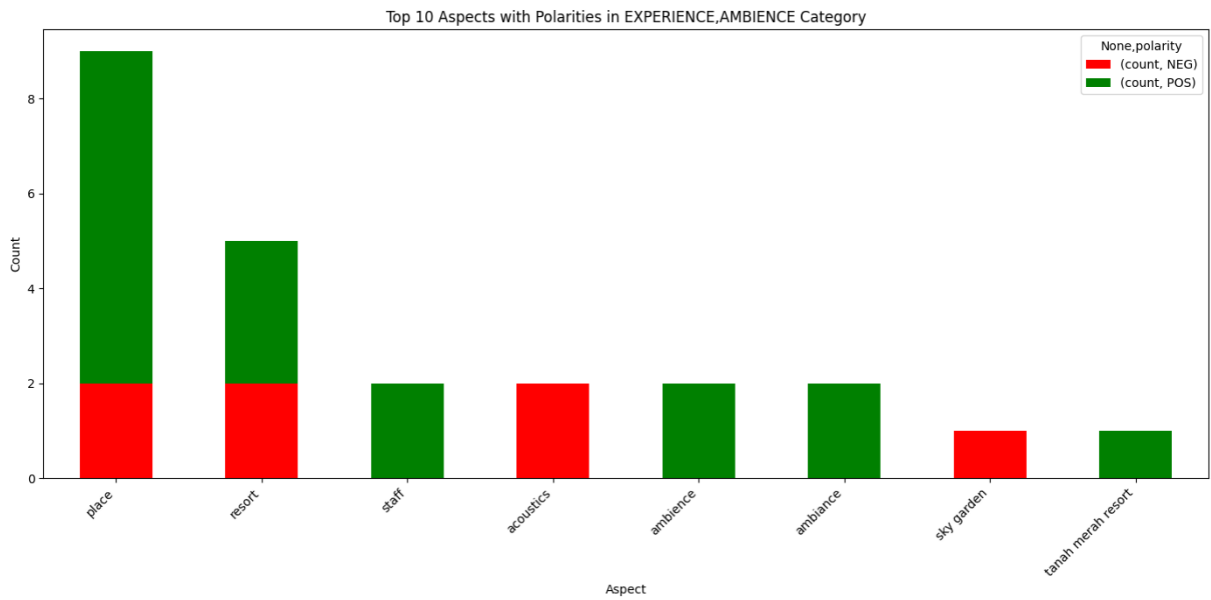
**Gambar 4.17 Sebaran Aspek Kategori Layanan**



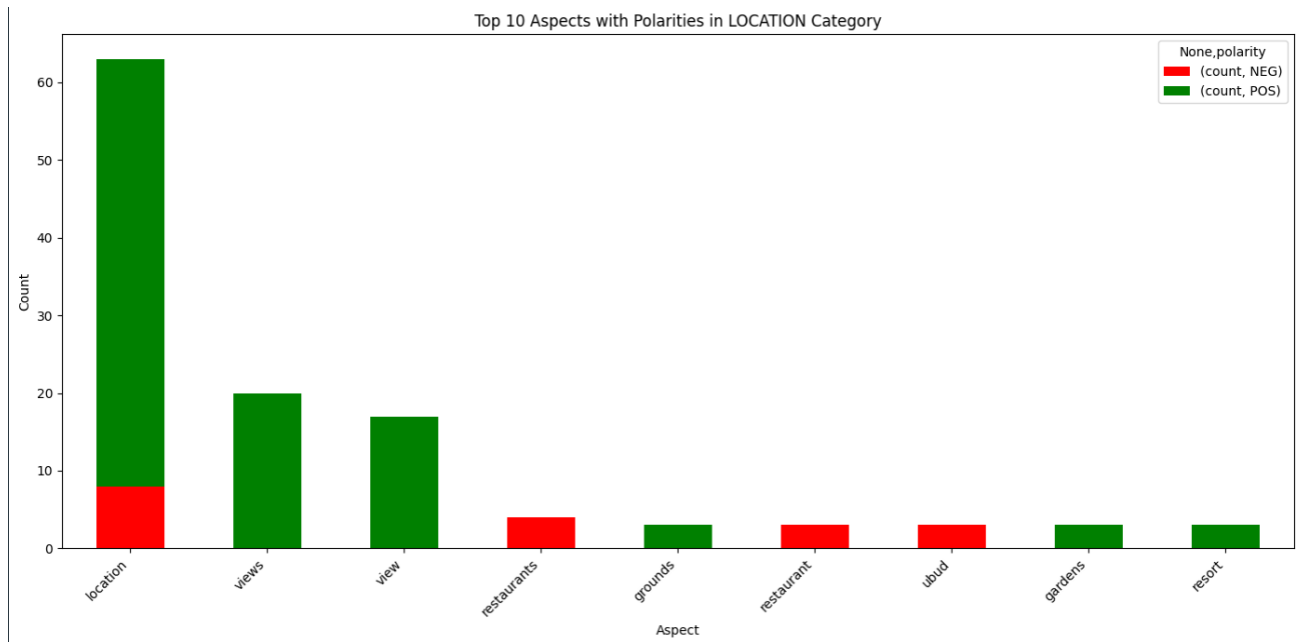
**Gambar 4.18 Sebaran Aspek Kategori Makanan**



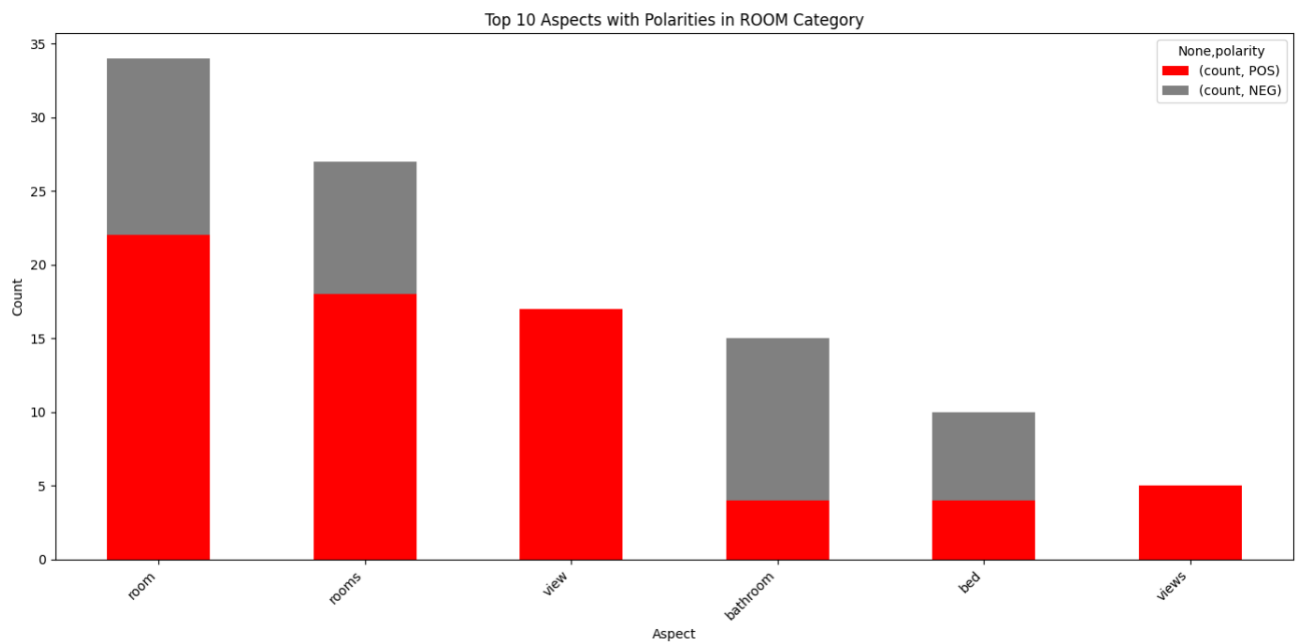
**Gambar 4.19** Sebaran Aspek Kategori Hotel



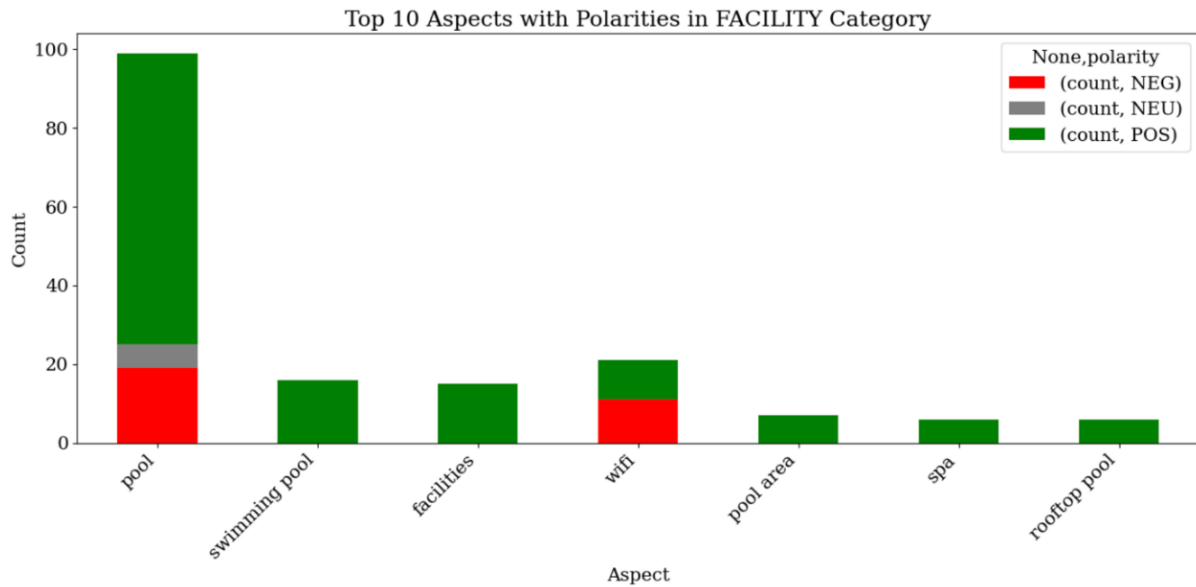
**Gambar 4.20** Sebaran Aspek Kategori



**Gambar 4.21 Sebaran Aspek Lokasi**



**Gambar 4.22 Sebaran Aspek Kamar**

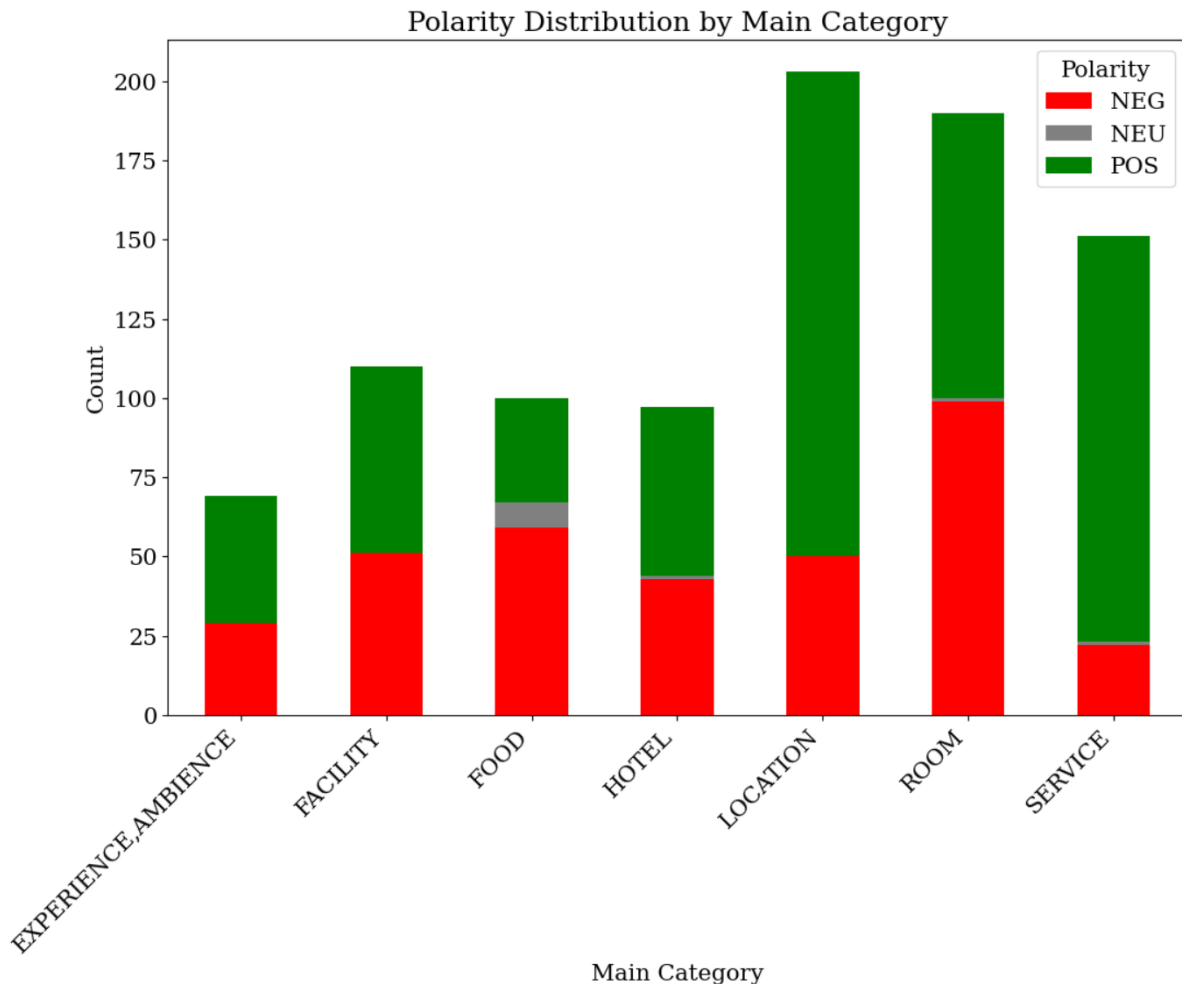


**Gambar 4.23 Sebaran Aspek Fasilitas**

Dari grafik grafik 4.17 – 4.23, terlihat aspek-aspek apa saja yang sering dibahas pada masing-masing kategori. Terdapat beberapa kategori yang hanya banyak membahas sedikit aspek, seperti kategori “FACILITY” mayoritas membahas “pool”, dan kategori “SERVICE” mayoritas membahas aspek “staff”. Pada beberapa kategori juga ada aspek yang cukup banyak dibahas seperti kategori kamar dan hotel, dengan kategori kamar banyak membahas aspek kamar mandi, *room*, *rooms*, dan pemandangan. Di sini aspek *room* dan *rooms* merupakan dua hal yang dapat dikelompokkan menjadi satu. Sedangkan kategori hotel banyak membahas aspek *property*, *hotel*, tempat, dan *resort*. Pada kategori hotel sendiri banyak aspek yang sebenarnya memiliki arti yang sama sehingga dapat dikelompokkan menjadi satu kategori. Dari

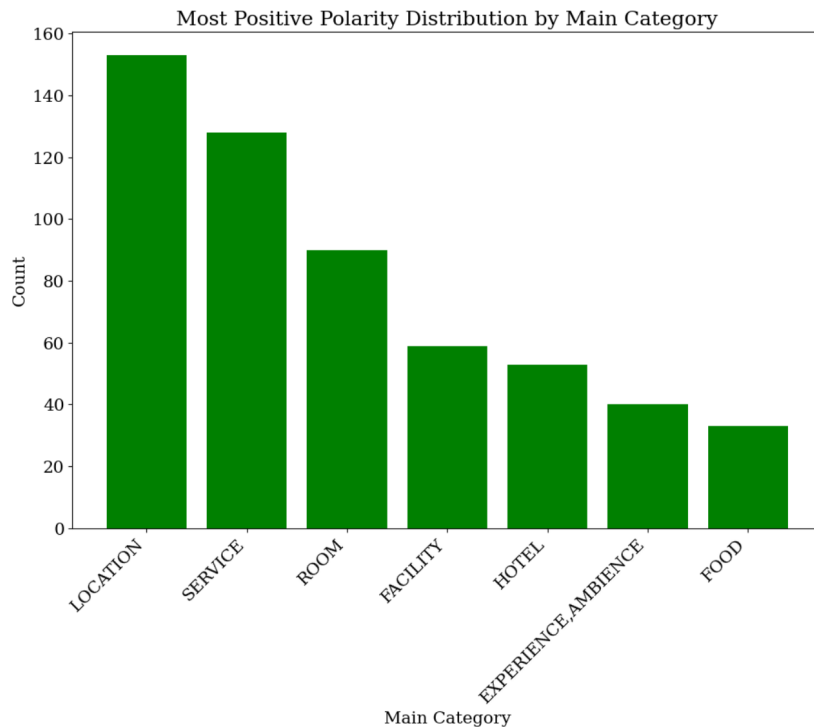
sini, dapat terlihat manfaat penggunaan kategori, yaitu mengelompokkan aspek aspek yang sekiranya dapat dikelompokkan sehingga dapat lebih mudah dilakukan analisis.

#### 4.6.3.4 Analisis Kategori



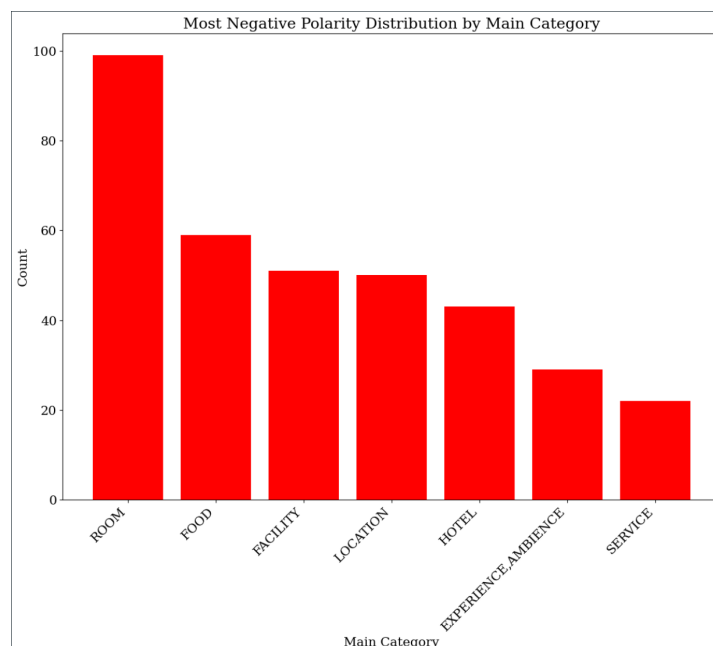
**Gambar 4.24 Distribusi Sentimen Kategori**

Grafik 4.24 menunjukkan bahwa kategori “LOCATION”, “ROOM”, dan “SERVICE” merupakan tiga kategori yang paling banyak dibahas. Ini cukup menarik karena customer sangat memperhatikan faktor lokasi pada hotel dengan kategori lokasi merupakan kategori yang paling banyak dibahas. Dari visualisasi ini, tidak cukup diketahui manakah kategori yang mempunyai banyak sentimen positif ataupun kategori yang memiliki banyak sentimen positif. Oleh karena itu, perlu dilakukan analisis tentang kedua sentimen secara terpisah.



**Gambar 4.25 Distribusi Sentimen Positif**

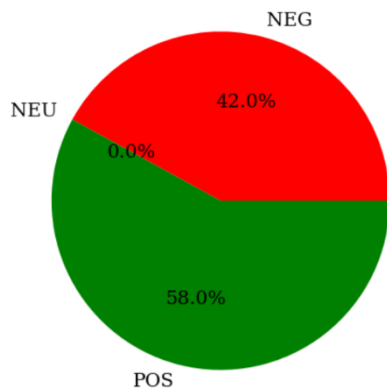
Gambar 4.25 menunjukkan urutan kategori berdasarkan sentimen terbanyak. Dapat disimpulkan ternyata kategori lokasi, pelayanan, dan kamar memiliki sentimen positif terbanyak.



**Gambar 4.26 Distribusi Sentimen Negatif**

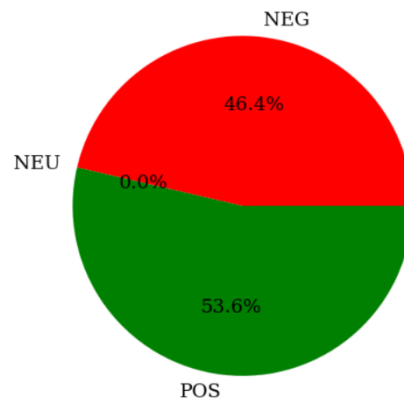
Grafik 4.26 menunjukkan bahwa kategori “ROOM” memiliki sentimen negatif terbanyak, hampir 100 sentimen negatif. Padahal kategori “ROOM” merupakan urutan ketiga sentimen positif terbanyak, yaitu melebihi 80 sentimen. Akan tetapi, ternyata lebih banyak sentimen negatif pada “ROOM” ketimbang sentimen positifnya. Sementara itu, kategori “FOOD”, “FACILITY”, “LOCATION”, “HOTEL” mempunyai sentimen negatif antara 40-60 sentimen.

Polarity Distribution for EXPERIENCE, AMBIENCE



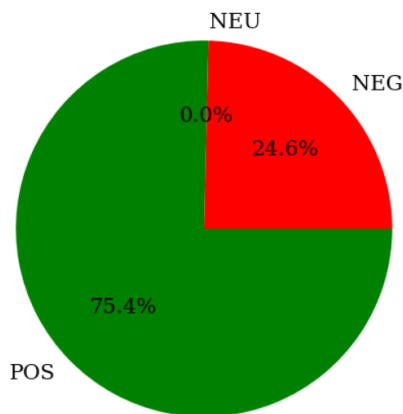
**Gambar 4.27 Distribusi Polaritas Kedua**

Polarity Distribution for FACILITY



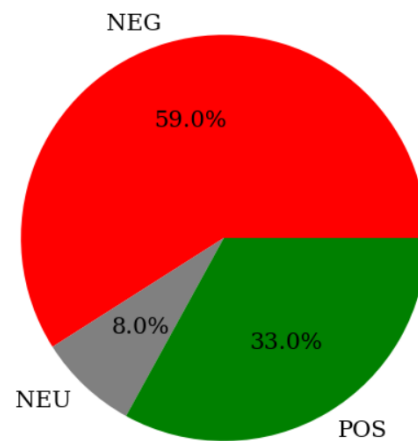
**Gambar 4.28 Distribusi Polaritas Fasilitas**

Polarity Distribution for LOCATION



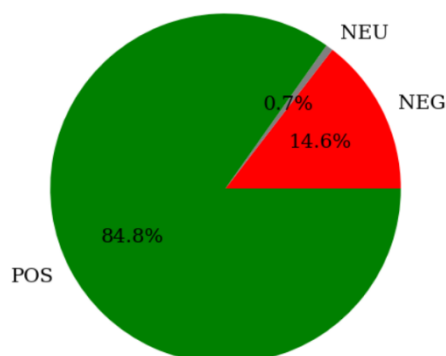
**Gambar 4.29 Distribusi Polaritas Kategori Lokasi**

Polarity Distribution for FOOD



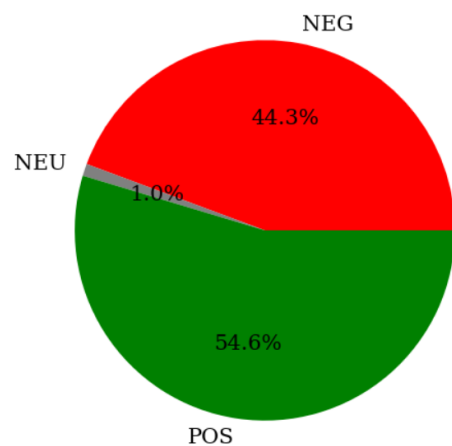
**Gambar 4.30 Distribusi Polaritas Kategori Makanan**

Polarity Distribution for SERVICE



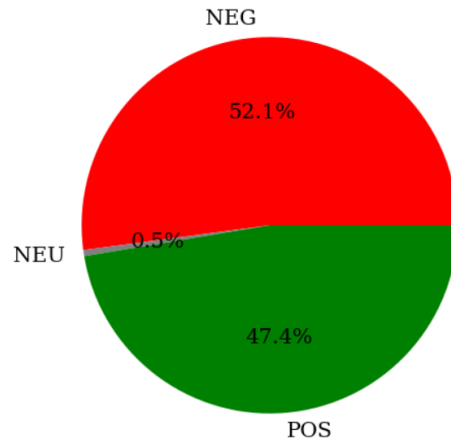
**Gambar 4.31 Distribusi Polaritas Kategori Layanan**

Polarity Distribution for HOTEL



**Gambar 4.32 Distribusi Polaritas Kategori Hotel**

Polarity Distribution for ROOM

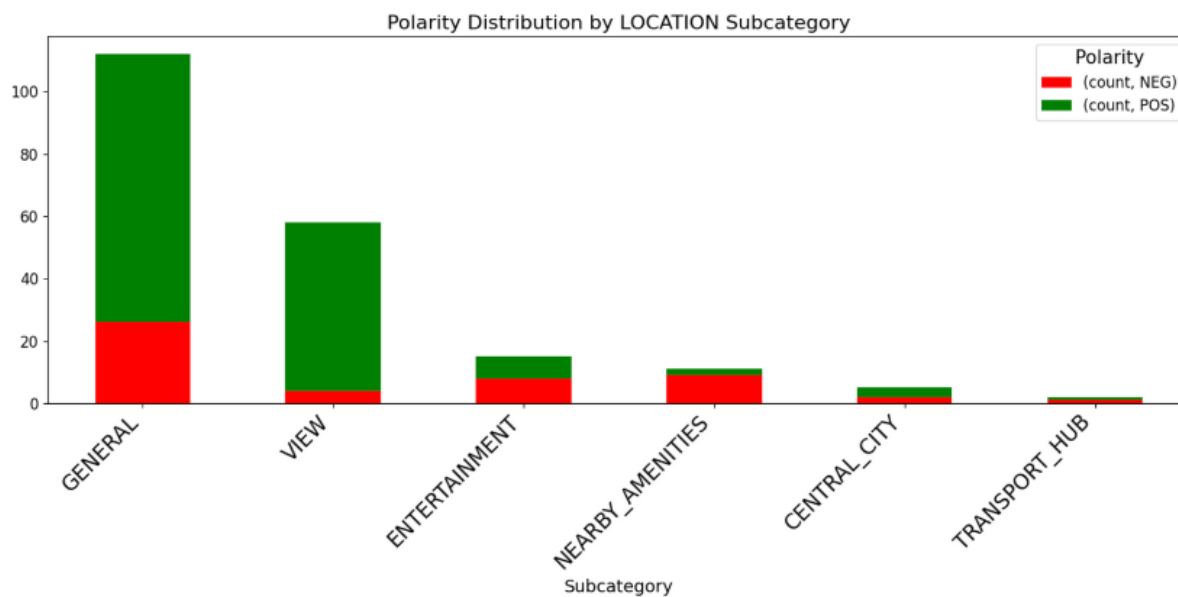


**Gambar 4.33 Distribusi Polaritas Kategori Kamar**

Pada gambar 4.27 – 4.33, terdapat visualisasi mengenai perbandingan sentimen positif, negatif, dan netral pada masing masing kategori. Kategori “ROOM”, “HOTEL”, “FACILITY”, dan “EXPERIENCE”, memiliki perbandingan sentimen dan negatif yang tidak terlalu jauh. Ini menunjukkan kategori-kategori tersebut masih banyak memiliki sentimen yang negatif, bahkan beberapa kategori memiliki sentimen negatif lebih banyak ketimbang sentimen positif. Kategori “FOOD” memiliki proporsi sentimen negatif terbanyak, yaitu 59% dari seluruh sentimen kategori tersebut. Sementara itu, kategori “SERVICE” dan “LOCATION” memiliki proporsi sentimen positif yang jauh lebih besar ketimbang sentimen negatifnya.

#### 4.6.3.5 Analisis kelebihan hotel

Berdasarkan analisis yang dilakukan sebelumnya, kategori lokasi dan pelayanan adalah dua kategori dengan sentimen positif terbanyak dan proporsi sentimen positif terbanyak.

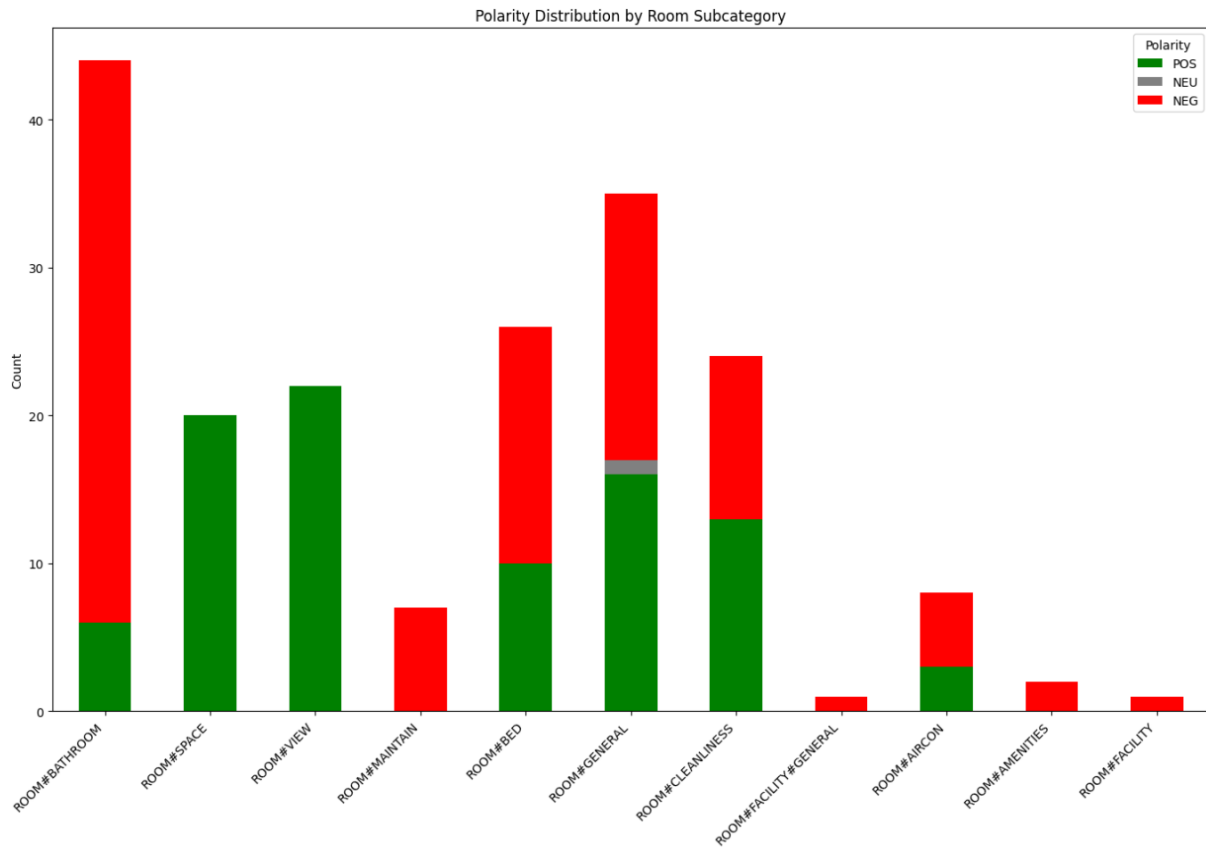


**Gambar 4.34 Distribusi Sub Kategori Lokasi**





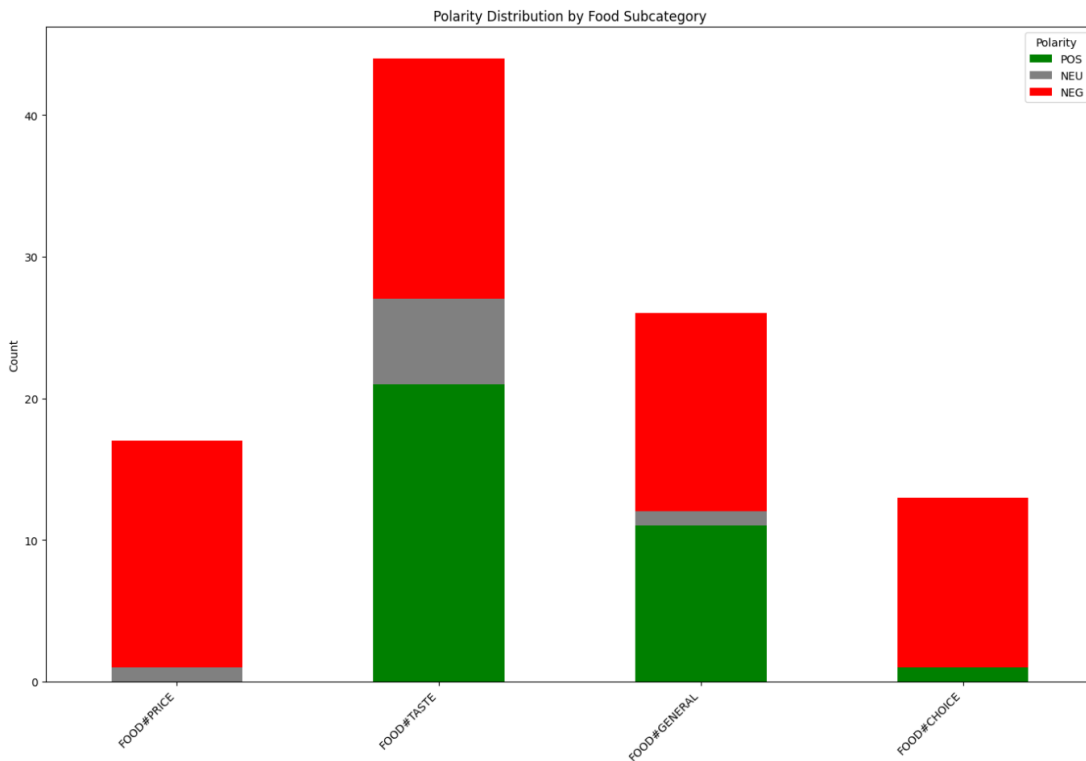




**Gambar 4.37 Bar-chart sub-kategori “ROOM”**

Berdasarkan gambar 4.37, banyak sekali sentimen negatif pada sub-kategori “ROOM#BATHROOM”, “ROOM#MANTAIN”, “ROOM#BED”, “ROOM#GENERAL”, “ROOM#CLEANLINESS” bahkan “ROOM#AIRCON”. Ini menunjukkan bahwa aspek aspek terkait kamar mandi, *maintanance* kamar, tempat tidur, kebersihan, serta *air conditioner* hotel ini mempunyai banyak kekurangan. Di samping itu, pada kategori “ROOM” luas kamar dan pemandangan, pengunjung mempunyai banyak sentimen positif.





**Gambar 4.39 Bar chart Distribusi Kategori “FOOD”**

Dari gambar 4.39, dapat dilihat bahwa banyak sentimen negatif pada “FOOD#PRICE”, “FOOD#TASTE”, “FOOD#GENERAL”, dan “FOOD#CHOICE”. Ini menunjukkan banyak pengunjung yang tidak cocok dengan harga makanan, rasa, serta pilihan-pilihan makanan yang diberikan oleh hotel.



**Gambar 4.40**

Gambar 4.40 menunjukkan kata-kata yang sering muncul pada opini kategori “FOOD”, di antaranya adalah *good*, *limited*, *expensive*, *poor*, *options*, *average*. Ini cukup mewakili kategori ini, *limited* mengarah sentimen negatif pada “FOOD#CHOICE” dan *expensive* mengarah sentimen negatif “FOOD#PRICE”. Pada tabel 4.18 ditunjukkan teks yang membahas tentang makanan bersama dengan *quadruple*-nya.

**Tabel 4.18 Contoh Teks yang Membahas “FOOD”**

Teks	"The hotel was not 5 star. 3 stars is a better rating for this hotel. <b>Restaurant food</b> was average. No <b>bartender</b> so a Bacardi cola was not possible. <b>Wine</b> list was very limited, with just 2 wines.
Quadruple	[restaurant food', 'NEU', 'average', 'FOOD#TASTE'], ['bartender', 'NEG', 'No', 'SERVICE#GENERAL'], ['Wine', 'NEG', 'limited', 'FOOD#CHOICE']

#### 4.6.3.7 Kesimpulan analisis hotel

Berdasarkan analisis yang dilakukan dengan memanfaatkan *quadruple* pada teks ulasan hotel, ditemukan *insight-insight* yang menarik tentang hotel :

1. Dari 902 *quadruple*, ulasan pada hotel menunjukkan 60.4% memiliki sentimen positif, sedangkan 38,4% nya memiliki sentimen negatif. Meskipun sentimen negatifnya lebih sedikit, tetapi ini merupakan angka dan proporsi yang cukup besar.
2. Hotel memiliki keunggulan di kategori lokasi dan kategori pelayanan. Kedua kategori memiliki sentimen positif sebesar kira kira 85% dan 75% Untuk kategori lokasi, banyak sentimen positif terhadap pemandangan dari hotel.
3. Hotel ini memiliki beberapa kekurangan, terutama di kategori kamar dan makanan. Kategori kamar dan makanan memiliki sentimen negatif berturut-turut sebesar kira-kira 59% dan 52%. Untuk kategori kamar, banyak pelanggan memiliki sentimen negatif pada kebersihan, kamar mandi, tempat tidur, *maintanance*, hingga pendingin udara. Sementara itu, pada kategori makanan, banyak sentimen negatif terhadap harga makanan, rasa makanan, serta variasi atau pilihan makanan.

## BAB 5 Kesimpulan dan Saran

### 5.1 Kesimpulan

Kesimpulan yang didapatkan dari pengerjaan tugas akhir yang telah dilakukan adalah:

1. Hasil *training* model menunjukkan perbedaan pada hasil evaluasi ketika dilatih dengan panjang teks yang berbeda. Saat dilatih menggunakan dataset *multi sentences* dan *single sentences* dengan parameter yang sama, model mempunyai nilai f1 sebesar 0.47 dan 0.17. Setelah dilakukan penyederhanaan kategori, Meskipun hasil metrik model dari dataset *multi sentences* cukup rendah, kenyataannya saat dianalisis, model cukup baik dalam melakukan prediksi. Akan tetapi, karena aspek yang ditangkap banyak seiring dengan panjangnya kalimat, jawaban model akan dianggap salah jika salah satu dari banyak jawabannya salah. Sementara model yang dilakukan pada dataset *single-sentence*, performa model mendekati hingga melebihi performa model *quadruple* pada *baseline model*.
2. Setelah melakukan penyederhanaan kategori pada kategori yang mempunyai arti yang mirip, model *single sentence* dapat mencapai nilai f1 sebesar 0.55. Performa model dalam memprediksi kategori pada teks memiliki nilai F1 sebesar 0.7553 dan pada prediksi sub-kategori memiliki nilai F1 sebesar 0.6331. Menentukan kategori pada *quadruple* bukan merupakan task yang sederhana karena kategori ditentukan tidak hanya pada aspek saja, tetapi juga opini dan konteks pada kalimat.
3. Model ASQE terbukti dapat menguraikan *quadruple* teks ulasan hotel dengan setiap *quadruple* terdiri dari aspek, sentimen, opini, dan kategori. Kumpulan *quadruple* ini dapat dimanfaatkan untuk keperluan analisis lebih mendalam.
4. Model berhasil diimplementasikan pada data sebuah hotel di Bali. Dari hasil prediksi *quadruple* data dibuktikan bahwa dapat menarik informasi dengan mudah yang dibantu dengan analisis dan visualisasi data, seperti melihat kelebihan dan kekurangan hotel, apa hal yang sering dibahas pada teks ulasan, dan opini yang sering muncul. Semua informasi ini dapat didapat tanpa melihat teks ulasan secara keseluruhan. Kategorisasi juga terbukti memudahkan data untuk dianalisis. Dengan adanya kategori, *quadruple* dapat dikelompokkan dengan kelompok yang jauh lebih sedikit ketimbang dikelompokkan dengan aspek yang sangat banyak.

### 5.2 Saran

Dari pengembangan tugas akhir ini, berikut adalah beberapa saran yang dapat diberikan untuk pengembangan penelitian di masa mendatang:

1. Melakukan penelitian metode -metode fine-tuning yang biasa dilakukan pada LLM (Large Language Model) seperti instruction tuning dan PEFT
2. Membuat pipeline yang memudahkan dalam penggunaan model, seperti otomatisasi dan *scheduler* menggunakan Apache Airflow.
3. Melakukan deployment pada model machine learning
4. Menerapkan *pipeline* model pada domain selain hotel.
5. Mengembangkan model yang dapat menganalisis teks dalam Bahasa Indonesia.





## DAFTAR PUSTAKA

Aliansyah, H., & Hermawan, W. (2021, February 26). PERAN SEKTOR PARIWISATA PADA PERTUMBUHAN EKONOMI KABUPATEN/KOTA DI JAWA BARAT. *Bina Ekonomi*, 23(1), 39–55. <https://doi.org/10.26593/be.v23i1.4654.39-55>  
Badan Pusat Statistik. (2021). Statistik Wisatawan Nusantara.

Cai H, Xia R, Yu J (2021) Aspect-category-opinion-sentiment quadruple extraction with implicit aspects and opinions. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pp 340–350

Fan, F., Feng, Y., & Zhao, D. (2018). Multigrained attention network for aspect-level sentiment classification. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 3433-3442

Fang, X., & Tao, J. (2019, October). A Transfer Learning based Approach for Aspect Based Sentiment Analysis. 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS). <https://doi.org/10.1109/snams.2019.8931817>

Hoang, C. D., Dinh, Q. V., & Tran, N. H. (2022, December 20). Aspect-Category-Opinion-Sentiment Extraction Using Generative Transformer Model. 2022 RIVF International Conference on Computing and Communication Technologies (RIVF). <https://doi.org/10.1109/rivf55975.2022.10013820>

Ma, D., Li, S., Zhang, X., & Wang, H. (2017). Interactive attention networks for aspect-level sentiment classification. arXiv preprint arXiv:1709.00893.

Liu, B. (2012). Sentiment Analysis and Opinion Mining (Issue April). Morgan & Claypool. [https://doi.org/10.1142/9789813100459\\_0007](https://doi.org/10.1142/9789813100459_0007).

Muhammad, P. F., Kusumaningrum, R., & Wibowo, A. (2021). Sentiment Analysis Using Word2vec and Long Short-Term Memory (LSTM) for Indonesian Hotel Reviews. *Procedia Computer Science*, 179, 728–735. <https://doi.org/10.1016/j.procs.2021.01.061>

Pusdatin Kemenparekraf. (2023, Sept 4). Statistik Kunjungan Wisatawan Mancanegara Bulan Juli 2023

Yang, H., Zeng, B., Yang, J., Song, Y., & Xu, R. (2021, January). A multi-task learning model for Chinese-oriented aspect polarity classification and aspect term extraction. *Neurocomputing*, 419, 344–356. <https://doi.org/10.1016/j.neucom.2020.08.001>

Yang, H., Zhang, C., & Li, K. (2022, August 2). PyABSA: A Modularized Framework for Reproducible Aspect-based Sentiment Analysis. arXiv.org. <https://arxiv.org/abs/2208.01368v3>

Zhang, H., Chen, Z., Chen, B., Hu, B., Li, M., Yang, C., & Jiang, B. (2022, July). Complete quadruple extraction using a two-stage neural model for aspect-based sentiment analysis. *Neurocomputing*, 492, 452–463. <https://doi.org/10.1016/j.neucom.2022.04.027>

Zeng, B., Yang, H., Xu, R., Zhou, W., & Han, X. (2019, August 17). LCF: A Local Context Focus Mechanism for Aspect-Based Sentiment Classification. MDPI. <https://doi.org/10.3390/app9163389>

BADAN PUSAT STATISTIK. (2022, December 27). STATISTIK HOTEL DAN AKOMODASI LAINNYA DI INDONESIA 2022. Retrieved October 11, 2023, from <https://www.bps.go.id/publication/2022/12/27/efc507a7343d35f9e0174421/statistik-hotel-dan-akomodasi-lainnya-di-indonesia-2022.html>

DataIndonesia.id. (2023, July 25). Sederet Destinasi Favorit Wisatawan Nusantara pada 2022 (Pratiwi, Ed.). Retrieved October 11, 2023, from <https://dataindonesia.id/sektor-riil/detail/sederet-destinasi-favorit-wisatawan-nusantara-pada-2022>

Ziaulhaq, Ibadurrahman (2023) Analisis Dan Kategorisasi Triplet Sentimen Level Aspek Pada Ulasan Hotel Menggunakan Span-Aste Dan Wu-Palmer Similarity Score (Studi Kasus: Provinsi Bali). Other thesis, Institut Teknologi Sepuluh Nopember.

Kusuma, Moch. Farrel Arrizal (2023) Pengembangan Model Analisis Sentimen Berbasis Aspek Dengan Teknik Transfer Learning Dan Local Sentiment Aggregation Pada Teks Bahasa (Studi Kasus : Ulasan Pariwisata Provinsi Bali). Other thesis, Institut Teknologi Sepuluh Nopember.

Apa itu Pemrosesan Bahasa Alami (NLP)? - Penjelasan tentang NLP - AWS. (n.d.). Amazon Web Services, Inc. <https://aws.amazon.com/id/what-is/nlp/>

Devlin, J., et al. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805.

Horev, R. (2018, November). BERT Explained: State of the art language model for NLP | by Rani Horev | Towards Data Science. <https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270>

Gupta, S. (2018). Sentiment Analysis: Concept, Analysis and Applications. <https://towardsdatascience.com/sentiment-analysis-concept-analysis-and-applications6c94d6f58c1>

Zhang, H., Cheah, Y. N., Alyasiri, O. M., & An, J. (2023, June 27). A Survey on Aspect-Based Sentiment Quadruple Extraction with Implicit Aspects and Opinions. Research Square (Research Square); Research Square (United States). <https://doi.org/10.21203/rs.3.rs-3098487/v1>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017, June 12). Attention Is All You Need. arXiv.org. <https://arxiv.org/abs/1706.03762v7>

He P., Liu X., Gao J., Chen W. (2020, June 5). DeBERTa: Decoding-enhanced BERT with Disentangled Attention. <https://arxiv.org/abs/2006.03654>

Nugroho, K. S. (2019, November). Confusion Matrix untuk Evaluasi Model pada Supervised Learning | by Kunchahyo Setyo Nugroho | Medium. <https://ksnugroho.medium.com/confusion-matrix-untuk-evaluasi-model-padaunsupervised-machine-learning-bc4b1ae9ae3f>

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2019, October 23). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. arXiv.org. <https://arxiv.org/abs/1910.10683v4>

[Hu M, Bai Y, Wu Y, et al \(2023\) Uncertainty-aware unlikelihood learning improves generative aspect sentiment quad prediction. arXiv preprint arXiv:230600418](#)

[Hu M, Wu Y, Gao H, et al \(2022\) Improving aspect sentiment quad prediction via template-order data augmentation. arXiv preprint arXiv:221010291](#)

[Bao X, Wang Z, Jiang X, et al \(2022\) Aspect-based sentiment analysis with opinion tree generation. IJCAI 2022 pp 4044–4050](#)

## LAMPIRAN

### Lampiran 1: Contoh Kelas ATE dan APC

```
class ATEInstruction(Instruction):
    def __init__(self, bos_instruction=None, eos_instruction=None):
        super().__init__(bos_instruction, eos_instruction)
        if self.bos_instruction is None:
            self.bos_instruction = f""
    Definition: The input are sentences about a product or service. The task is to extract the aspects.
    Here are some examples:

    example 1-
    input: I charge it at night and skip taking the cord with me because of the good battery life.
    {self.eos_instruction}
    aspect:battery life|aspect:cord

    example 2-
    input: Great food, good size menu, great service and an unpretentious setting.
    {self.eos_instruction}
    aspect:food|aspect:menu|aspect:service|aspect:setting

    Now extract aspects from the following example:
    input: """

        if self.eos_instruction is None:
            self.eos_instruction = "\nlet us extract aspects one by one: \n"

        if not self.bos_instruction:
            self.bos_instruction = bos_instruction
        if not self.eos_instruction:
            self.eos_instruction = eos_instruction

    def prepare_input(self, input_text):
        return self.bos_instruction + input_text + self.eos_instruction

class APCInstruction(Instruction):
    def __init__(self, bos_instruction=None, eos_instruction=None):
        super().__init__(bos_instruction, eos_instruction)
        if self.bos_instruction is None:
            self.bos_instruction = f""
    Definition: The input are sentences about a product or service. The task is to extract the aspects and
    their corresponding polarity. Here are some examples:

    example 1-
    input: I charge it at night and skip taking the cord with me because of the good battery life.
    The aspects are: battery life, cord
    {self.eos_instruction}
    battery life:positive|cord:positive

    example 2-
    input: Great food, good size menu, great service and an unpretentious setting.
    The aspects are: food, menu, service, setting
    {self.eos_instruction}
    food:positive|menu:positive|service:positive|setting:positive

    Now predict aspect sentiments from the following example:
    input: """

        if self.eos_instruction is None:
            self.eos_instruction = "\nlet us predict sentiments one by one: \n"

        if not self.bos_instruction:
            self.bos_instruction = bos_instruction
        if not self.eos_instruction:
            self.eos_instruction = eos_instruction

    def prepare_input(self, input_text, aspects):
        return (
            self.bos_instruction
            + input_text
            + f"The aspects are: {aspects}"
            + self.eos_instruction
        )
```

## **Lampiran 2: Link Github Repository**

<https://github.com/kwang0149/ta-final-project-aspect-based-sentiment-analysis-and-categorized-aspect>

## BIODATA PENULIS



Penulis lahir di Surabaya, 19 Februari 2002. Penulis menempuh pendidikan di SDK Santa Theresia Surabaya, lalu melanjutkan pendidikan di SMP Negeri 1 Surabaya, kemudian dilanjutkan di SMA Negeri 5 Surabaya. Setelah menamatkan SMA, penulis melanjutkan pendidikan di perguruan tinggi Institut Teknologi Sepuluh Nopember Surabaya dengan NRP 5026201127.

Pada masa perkuliahan, penulis aktif mengikuti kegiatan di dalam perkuliahan dan di luar perkuliahan, seperti mengikuti studi independen Bangkit *machine learning path* by Google, Traveloka, menjadi asisten dosen pada mata kuliah Algoritma dan Struktur Data, serta mengikuti magang dengan role Backend Developer di Telkom.

Di akhir perkuliahan, penulis melakukan penelitian di laboratorium Rekayasa Data dan Intelegensi Bisnis departemen Sistem Informasi ITS. Jika ada yang ingin didiskusikan, ditanyakan mengenai penelitian ini maupun menghubungi penulis, penulis dengan senang hati menerimanya dan dapat dihubungi di [hongkwang321@gmail.com](mailto:hongkwang321@gmail.com)