



Identifikasi Multi Aspek dan Sentimen Analisis pada Review Hotel Menggunakan *Deep learning*

Windi Astriningsih

19917017

Tesis diajukan sebagai syarat untuk meraih gelar Magister Komputer

Konsentrasi Sains Data

Program Studi Informatika Program Magister

Fakultas Teknologi Industri

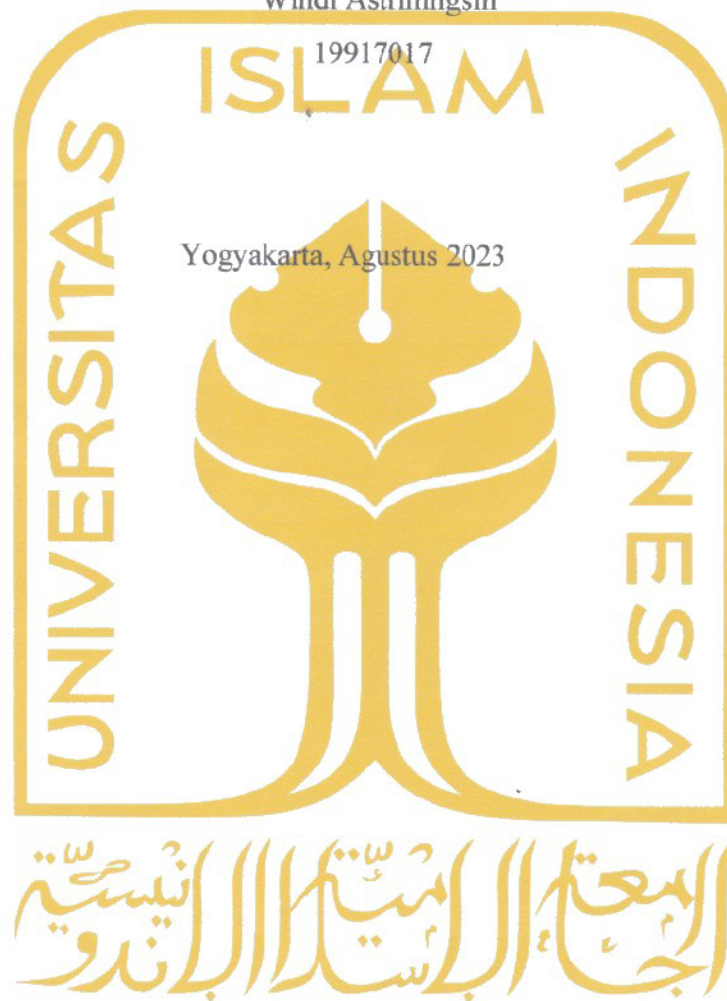
Universitas Islam Indonesia

2023

Lembar Pengesahan Pembimbing

Multi Aspek Sentimen Analisis pada Review Hotel Menggunakan *Deep Learning*

Windi Astriningsih



Pembimbing 1

Pembimbing 2


Dhomas Hatta Fudholi, S.T., M.Eng., Ph.D.


Chandra Kusuma Dewa, S.Kom., M.Cs.,
Ph.D.

Lembar Pengesahan Penguji

**Identifikasi Multi Aspek dan Sentimen Analisis pada Review Hotel
Menggunakan *Deep learning***

Windi Astriningsih

19917017

ISLAM

Yogyakarta, Agustus 2023

Tim Penguji,

Dhomas Hatta Fudholi, S.T., M.Eng., Ph.D.

Ketua

Chandra Kusuma Dewa, S.Kom., M.Cs., Ph.D.

Anggota I

Dr. Syarif Hidayat, S.Kom., M.I.T.

Anggota II

Mengetahui,

Ketua Program Studi Teknik Informatika Program Magister

Universitas Islam Indonesia



Irving Vitra Paputungan, S.T., M.Sc., Ph.D.

Abstrak

Identifikasi Multi Aspek dan Sentimen Analisis pada Review Hotel Menggunakan *Deep learning*

Review hotel tidak hanya memberikan informasi bermanfaat bagi pengusaha, tetapi juga membentuk citra hotel di mata pelanggan. Review umumnya mencakup berbagai aspek yang diungkapkan secara jujur oleh pelanggan. Dalam penelitian terdahulu (Cahyaningtyas et al., 2021; Jayanto et al., 2022; Tran et al., 2019) telah dikembangkan model berbasis aspek dan sentimen review untuk mengetahui penilaian dari sebuah review hotel. Namun dalam penelitian sebelumnya, masih melakukan identifikasi terhadap satu nilai tertentu dalam rangkaian kalimat. Sementara itu, pada kebanyakan review umumnya terdapat lebih dari satu penilaian. Oleh karena itu, untuk melihat berbagai penilaian review hotel dikembangkan analisis sentimen berbasis multi-aspek. Pengembangan ini bertujuan untuk mengidentifikasi lebih dari satu aspek dalam sebuah kalimat review dan sentimen. Aspek penilaian pada review diperoleh dari penelitian (Cahyaningtyas et al., 2021) yang meliputi aspek harga, lokasi, pelayanan, makanan, fasilitas, dan kamar. Menggunakan metode deep learning yaitu LSTM diharapkan model mampu secara efektif menangani rangkaian kata yang panjang. Arsitektur model LSTM terdiri dari layer embedding, SpatialDropout1D, LSTM, dan dense. Model dilatih dengan pengaturan epochs sebesar 10 dan batch size 32. Evaluasi model dilakukan dengan tiga skenario yang mencakup pengujian kalimat satu aspek, pengujian kalimat kombinasi dua aspek, dan pengujian kalimat kombinasi tiga aspek. Akurasi kombinasi K-aspek diterapkan pada pengujian kalimat kombinasi dua dan tiga aspek. Sementara F1_score digunakan untuk pengujian kalimat satu aspek dan sentimen kalimat. Hasil pengujian yang diperoleh yaitu akurasi kombinasi 2-aspek sebesar 79% pada kalimat kombinasi dua aspek dan tiga aspek, F1_score sebesar 85,7% pada kalimat satu aspek, dan F1_score sebesar 83% pada sentiment kalimat. Perolehan hasil tersebut menunjukkan bahwa model yang dibangun sudah mampu melakukan sentiment multi aspek pada review hotel.

Kata kunci

Sentimen Analysis, LSTM, Review Hotel, Multi Aspek

Abstract

Identifying Multi-Aspect and Sentiment Analysis in Hotel Reviews Using Deep Learning

Hotel reviews not only provide valuable information for business owners but also shape the image of the hotel in the eyes of customers. Reviews generally cover various aspects expressed honestly by customers. In previous studies (Jayanto et al., 2022; Priyantina & Sarno, 2019; Tran et al., 2019), aspect-based and sentiment-based review models were developed to understand the assessment of a hotel review. However, in previous research, identification was still carried out on a specific value within a sentence. Meanwhile, in most reviews, there is usually more than one assessment. Therefore, to examine various assessments of hotel reviews, a multi-aspect sentiment analysis was developed. This development aims to identify more than one aspect in a review sentence and the sentiment of the review. The assessment aspects in the review were obtained from research (Cahyaningtyas et al., 2021), which include aspects like price, location, service, food, facilities, and rooms. Using deep learning methods such as LSTM, it is expected that the model can effectively handle long sequences of words. The LSTM model architecture consists of embedding layers, SpatialDropout1D, LSTM, and dense layers. The model was trained with 10 epochs and a batch size of 32. Model evaluation was conducted with three scenarios, including testing sentences with one aspect, testing sentences with a combination of two aspects, and testing sentences with a combination of three aspects. The accuracy of the K-aspect combinations was applied to testing sentences with a combination of two and three aspects. Meanwhile, F1_score was used for testing sentences with one aspect and sentence sentiment. The test results obtained are as follows: an accuracy of 79% for 2-aspect combinations in sentences with a combination of two and three aspects, an F1_score of 85.7% for sentences with one aspect, and an F1_score of 83% for sentence sentiment. These results indicate that the constructed model is capable of performing multi-aspect sentiment analysis on hotel reviews.

Keywords

Sentiment Analysis, LSTM, Hotel Review, Multi Aspect

Pernyataan Keaslian Tulisan

Dengan ini saya menyatakan bahwa tesis ini merupakan tulisan asli dari penulis, dan tidak berisi material yang telah diterbitkan sebelumnya atau tulisan dari penulis lain terkecuali referensi atas material tersebut telah disebutkan dalam tesis. Apabila ada kontribusi dari penulis lain dalam tesis ini, maka penulis lain tersebut secara eksplisit telah disebutkan dalam tesis ini.

Dengan ini saya juga menyatakan bahwa segala kontribusi dari pihak lain terhadap tesis ini, termasuk bantuan analisis statistik, desain survei, analisis data, prosedur teknis yang bersifat signifikan, dan segala bentuk aktivitas penelitian yang dipergunakan atau dilaporkan dalam tesis ini telah secara eksplisit disebutkan dalam tesis ini.

Segala bentuk hak cipta yang terdapat dalam material dokumen tesis ini berada dalam kepemilikan pemilik hak cipta masing-masing. Apabila dibutuhkan, penulis juga telah mendapatkan izin dari pemilik hak cipta untuk menggunakan ulang materialnya dalam tesis ini.

Yogyakarta, Agustus 2023

A handwritten signature in black ink is written over a yellow 10,000 Indonesian postage stamp. The stamp features the Garuda Pancasila emblem and the text '10000 METERAI TEMPEL' and 'FABAIX591930663'.

Windi Astriningsih, S.Kom

Daftar Publikasi

Publikasi yang menjadi bagian dari tesis

Astriningsih, Windi and Fudholi, Dhomas Hatta. *Identifikasi Multi Aspek dan Sentimen Analisis pada Review Hotel Menggunakan Deep Learning*. s.l. : JATISI (Jurnal Teknik Informatika dan Sistem Informasi) Volume 10 Nomor 3, 2023.

Sitasi publikasi 1

Kontributor	Jenis Kontribusi
Windi Astriningsih	Melakukan komputasi dan analisis model Menulis paper
Dhomas Hatta Fudholi	Melakukan analisis model Melakukan pengecekan dan memberikan masukan pada penulisan paper

Halaman Kontribusi

Dalam penulisan tesis ini pembimbing I dan II memberikan beberapa masukan sebagai perbaikan dari cara penulisan tesis serta memberikan saran tentang data yang akan diolah, dan dianalisis.

Halaman Persembahan

Tesis ini saya dedikasikan dengan setulus hati & sepenuh jiwa untuk :

Orang tuaku tercinta (Ibuku Endang Suwartijah dan Ayahku Samingun) yang selalu mengiringi langkah kakiku dengan do'a yang tak pernah putus.

Adik-adiku yang selalu memberikan support dan doa.

Seluruh rekan kerja yang selalu memberikan pengertian dan supportnya.

Dan untuk diri sendiri, terimakasih karena telah berjuang menyelesaikan tesis ini dengan maksimal.

Kata Pengantar

Alhamdulillahirabbil'alamin, segala puji syukur kehadiran Allah SWT senantiasa penulis panjatkan karena atas limpahan rahmat dan hidayah-Nya, tesis yang berjudul “Multi Aspek Sentimen Analisis pada Review Hotel Menggunakan Deep Learning” dapat berjalan dengan lancar dalam penyelesaiannya. Tesis ini diajukan sebagai bagian dalam menyelesaikan studi dan sebagai salah satu syarat untuk memperoleh gelar Magister Komputer pada Program Studi Magister Teknik Informatika Fakultas Teknologi Industri Universitas Islam Indonesia.

Dalam penyelesaian Tesis ini, penulis banyak mendapatkan bantuan dari berbagai pihak, untuk itu penulis menyampaikan ucapan terima kasih setulusnya kepada:

1. Bapak Prof. Fathul Wahid, S.T., M.Sc., Ph.D., selaku Rektor Universitas Islam Indonesia.
2. Bapak Irving Vitra Papatungan, S.T., M.Sc., Ph.D, selaku Ketua Program Studi Teknik Informatika Program Studi Magister Universitas Islam Indonesia.
3. Bapak DThomas Hatta Fudholi, S.T., M.Eng., Ph.D., selaku dosen pembimbing pertama, yang telah banyak membantu penulis dalam memberikan ide, saran dan kritiknya.
4. Bapak Chandra Kusuma Dewa, S.Kom., M.Cs., Ph.D, selaku dosen pembimbing kedua, yang membantu penulis dalam memberikan saran penelitian tesis.
5. Bapak/Ibu “penguji”., selaku dewan penguji tesis yang telah memberikan banyak pengarahan dan masukan dalam penyusunan dan penyempurnaan Tesis ini.
6. Seluruh Dosen Program Studi Magister Teknik Informatika yang telah memberikan bekal ilmu pengetahuan kepada penulis, semoga ilmunya menjadi amal jariyah di dunia maupun akhirat.
7. Staff Akademik Program Pascasarjana Fakultas Teknologi Universitas Islam Indonesia, yang telah membantu dalam segala urusan administrasi di kampus.
8. Rekan kerja di tim IT PT AVO Innovation Technology yang sudah cukup pengertian dengan jadwal kuliah saya.

Penulis menyadari bahwa dalam penulisan tesis ini masih banyak kelemahan dan kekurangan, untuk itu kritik dan saran yang sifatnya membangun sangat penulis harapkan agar tesis ini dapat menjadi lebih baik.

Yogyakarta, 2023

(Windi Astriningsih)

Daftar Isi

Lembar Pengesahan Pembimbing	ii
Lembar Pengesahan Penguji.....	Error! Bookmark not defined.
Abstrak	iii
Abstract.....	v
Pernyataan Keaslian Tulisan	vi
Daftar Publikasi	vii
Halaman Kontribusi.....	vii
Halaman Persembahan	ix
Kata Pengantar.....	x
Daftar Isi.....	x
Daftar Tabel.....	xii
Daftar Gambar	xiv
Glosarium	xvi
BAB 1 Pendahuluan	1
1.1 Pendahuluan.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB 2 Tinjauan Pustaka	6
2.1 Penelitian Terdahulu	6
2.2 Landasan Teori.....	10
2.2.1 Deep Learning	10
2.2.2 LSTM	11
2.2.3 Pengaturan Hyperparameter	13

2.2.4	<i>Preprocessing Data</i>	14
2.2.5	Metode pengukuran	15
BAB 3	Metode Penelitian.....	19
3.1	<i>Data Collection</i>	19
3.2	<i>Preprocessing Data</i>	21
3.3	Pembangunan Model Data.....	22
3.3.1	Flow Training Data.....	22
3.3.2	Arsitektur Training Data Review	25
3.3.3	Learning Curve Training Data.....	26
3.4	<i>Prediction Evaluation</i>	28
BAB 4	Hasil dan Pembahasan.....	31
4.1	Proses Pembangunan Model Multi Aspek.....	31
4.2	Proses Evaluasi Model dengan Kalimat Satu Aspek	38
4.3	Proses Evaluasi Model dengan Kalimat Multi Aspek	40
4.4	Proses Evaluasi Model Sentimen Aspek.....	43
4.5	Pembahasan	44
4.5.1	Hasil Akurasi Training Model.....	44
4.5.2	Hasil Evaluasi Klasifikasi Aspek	47
4.5.3	Hasil Evaluasi Sentimen.....	56
BAB 5	Kesimpulan dan Saran.....	58
5.1	Kesimpulan	58
5.2	Saran	58
	Daftar Pustaka	60
	LAMPIRAN	63

Daftar Tabel

Tabel 2.1 Penelitian Terkait <i>Deep Learning</i> Pada Analisis Data Teks	8
Tabel 3.1 Contoh Evaluasi Kombinasi K-aspek.....	29
Tabel 4.1 Jumlah Data Review Pada Masing-Masing Aspek.....	31
Tabel 4.2 Hasil <i>Training</i> Model Aspek dengan Berbagai LSTM Unit	45
Tabel 4.3 Learning Curve Pada Setiap Training Model Aspek dengan LSTM Unit yang Berbeda.....	46
Tabel 4.4 Hasil Evaluasi Kalimat Satu Aspek.....	48
Tabel 4.5 Hasil Evaluasi Kalimat Kombinasi Dua Aspek.....	52
Tabel 4.6 Hasil Evaluasi Kalimat Kombinasi Tiga Aspek	55

Daftar Gambar

Gambar 2.1 LSTM arsitektur.....	12
Gambar 2.2 Legenda notasi pada arsitektur LSTM.....	12
Gambar 2.3 Contoh confusion matrix pada model sentimen.....	17
Gambar 3.1 Gambaran alur penelitian.....	19
Gambar 3.2 Kumpulan kalimat review dengan aspek utama yaitu kamar.	20
Gambar 3.3 Contoh hasil <i>cleansing</i> dan <i>lowercase</i> pada kalimat review	21
Gambar 3.4 Contoh representasi kata dalam bentuk numerik.....	22
Gambar 3.5 Proses training model aspek.	23
Gambar 3.6 Proses training model sentiment.....	24
Gambar 3.7 Arsitektur model LSTM	25
Gambar 3.8 Learning curve model aspek makanan.....	27
Gambar 3.9 Confusion matrix pada analisis sentimen	30
Gambar 4.1 <i>Import library</i> training data	32
Gambar 4.2 <i>Import data training set</i>	33
Gambar 4.3 Melakukan tokenisasi pada data teks.....	33
Gambar 4.4 Mengambil aspek pada <i>training</i> data.....	34
Gambar 4.5 Melakukan proses pemodelan data	35
Gambar 4.6 Proses visualisasi learning curve training model.....	37
Gambar 4.7 Menyimpan model data yang telah dilakukan.	37
Gambar 4.8 <i>Import library</i> untuk proses validasi.....	38
Gambar 4.9 Inisiasi data evaluasi	39
Gambar 4.10 Melakukan tokenisasi pada data evaluasi kalimat satu aspek.....	39
Gambar 4.11 Melakukan evaluasi pada model dengan data teks	39
Gambar 4.12 Melakukan konfusi matriks pada pengujian kalimat satu aspek.....	40
Gambar 4.13 Melakukan perhitungan F1_Score pada hasil prediksi	40
Gambar 4.14 Melakukan evaluasi dua aspek.	41
Gambar 4.15 Melakukan perhitungan akurasi pada dua aspek.	42
Gambar 4.16 Melakukan evaluasi kalimat kombinasi tiga aspek.....	42
Gambar 4.17 Perhitungan <i>accuracy</i> pada kalimat kombinasi tiga aspek.	43
Gambar 4.18 Inisiasi data validasi sentiment	43
Gambar 4.19 Tahapan evaluasi sentimen.	44
Gambar 4.20 Evaluasi hasil sentimen dengan F1_Score.....	44

Gambar 4.21 Confusion matrix aspek makanan.....	49
Gambar 4.22 Confusion matrix aspek lokasi.....	49
Gambar 4.23 Confusion matrix aspek kamar	50
Gambar 4.24 Confusion matrix aspek pelayanan.....	50
Gambar 4.25 Confusion matrix aspek harga	51
Gambar 4.26 Hasil prediksi kalimat dengan aspek makanan.....	51
Gambar 4.27 Hasil prediksi model dengan kombinasi kalimat dua aspek.....	53
Gambar 4.28 Matriks hasil evaluasi sentiment kalimat.....	57
Gambar 4.29 Hasil prediksi sentimen review.....	57

Glosarium

LSTM	- Long Short Term Memory
TF	- Tensorflow
AI	- Artificial Intelligence

BAB 1

Pendahuluan

1.1 Pendahuluan

Review hotel menjadi salah satu dataset yang sangat penting untuk diteliti karena review terkait karakteristik di setiap hotel selalu berbeda (Tran et al., 2019). Review tersebut menjadi sumber informasi bagi para wisatawan untuk memilih tempat singgah saat dalam perjalanan. Selain itu, review juga menjadi sarana pengembangan kualitas hotel oleh pemilik usaha. Oleh karena itu, data review hotel dapat diolah untuk mengetahui penilaian yang terkandung dalam review dan memberikan informasi pengembangan industri perhotelan.

Sentimen analisis atau disebut juga *opinion mining* adalah bidang studi untuk menganalisa opini, sentimen, sikap, dan emosi manusia terhadap entitas yang melekat pada suatu barang, jasa, individu, atau topik tertentu (Liu et al., 2019). Sentimen analisis berbasis aspek terdiri dari ekstraksi aspek dan sentimen, serta penentuan terhadap orientasi suatu aspek. Orientasi dari aspek yang ditemukan dapat berupa positif, negatif, dan netral.

Penelitian terkait sentimen analysis berbasis aspek menggunakan data review hotel dilakukan oleh (Cahyaningtyas et al., 2021). Terdapat enam aspek pada penelitian tersebut yaitu harga, lokasi, pelayanan, kamar, hotel, dan restoran. Proses pengujian dilakukan dengan menerapkan beberapa metode *deep learning* diantaranya LSTM, RNN, CNN, Attention, dan GRU. Pada penelitian tersebut diperoleh hasil akurasi sebesar 0.926064 untuk proses klasifikasi aspek menggunakan metode LSTM. Namun masih ditemukan beberapa ketidaksesuaian klasifikasi ketika terdapat lebih dari satu aspek dalam satu dokumen opini.

Penelitian sentimen analysis berbasis aspek lainnya dilakukan oleh (Hoang & Rouces, 2019), (Yutika et al., 2021), (Mubarok et al., 2017). Dimana masing-masing melakukan penelitian pada data teks untuk mengetahui sentimen dari setiap aspek dalam level *document*. Menggunakan berbagai *scenario deep learning* pada ketiga penelitian tersebut diperoleh hasil akurasi yang cukup baik terhadap sentiment aspek. Sehingga menunjukkan bahwa dengan metode *deep learning* dapat diketahui polaritas suatu opini terhadap spesifik aspek didalamnya.

Selanjutnya penelitian berbasis aspek dikembangkan kedalam tahapan yang lebih kompleks yaitu sentimen analisis berbasis multi aspek. Pada penelitian (Ma et al., 2019), mencoba melakukan pemodelan multi aspek dalam satu opini yang tercantum pada sebuah

dokumen. Namun pada penelitian tersebut masih belum dilakukan proses untuk mengetahui kecenderungan sentimen dari opininya.

Sementara itu, penelitian (Hu et al., 2018) menunjukkan multi-aspek sentimen analisis dapat dilakukan dengan mengetahui posisi aspek tertentu dalam sebuah kalimat dengan metode *constrain attention network* (CAN). Metode yang digunakan dapat mengetahui letak dari aspek dalam kalimat dan kata yang menunjukkan opini terhadap aspek tersebut. Seperti contoh pada kalimat *'sometimes i get good food and ok service'*, metode CAN diterapkan untuk bisa melihat posisi di kiri dan kanan dari kata *food* dan *service*. Hal tersebut bertujuan untuk melihat aspek dan polarity pada rangkaian kalimat. Dalam proses pengujian model, terdapat dua kasus yang dicoba yaitu overlapping dan error. Kasus overlapping aspek dapat diidentifikasi dengan benar, dimana overlapping terjadi ketika terdapat dua aspek dalam kalimat dengan satu sentimen yang sama. Contoh kalimat overlapping yang diidentifikasi dengan benar *'I was highly disappointed by their service and food'*, dimana terdapat dua aspek yaitu *service* dan *food* dengan satu kata polarity yaitu *disappointed*. Sementara itu, error case masih terjadi ketika terdapat dua aspek yang berbeda dengan polarity yang berbeda pula seperti kalimat *'But dinner here is never disappointing, even if the prices are a bit over the top'*.

Beberapa penelitian terkait sentiment analysis berbasis aspek menunjukkan hasil akurasi yang cukup baik. Namun pengembangan penelitian terkait hal tersebut perlu dilakukan. Hal ini dijelaskan oleh (Bouazizi & Ohtsuki, 2017), dimana sentimen analysis berbasis aspek dianggap menilai secara keseluruhan sentimen yang ada pada suatu kalimat. Padahal bisa jadi terdapat beberapa aspek yang bisa diidentifikasi untuk menentukan maksud dari penulisan kalimat tersebut oleh manusia. Maka dalam hal ini, penelitian terkait sentimen berbasis aspek multi-kelas akan memberikan informasi yang lebih mendalam terhadap maksud penulisan sebuah kalimat.

Long-short Term Memory (LSTM) merupakan metode klasifikasi *deep learning* yang telah banyak digunakan pada penelitian beberapa tahun belakngandalam penelitian terkait sentiment analysis berbasis aspek menggunakan metode LSTM yang dikombinasikan dengan *Word Embedding* dilakukan oleh (Priyantina & Sarno, 2019). Hasil penelitian dengan kombinasi tersebut memperoleh akurasi klasifikasi sebesar 93%.

Dari hasil beberapa penelitian terdahulu, disimpulkan bahwa penilaian suatu kalimat tidak dapat diandalkan hanya dengan satu aspek saja. Hal ini menjadi penting dalam penulisan review hotel karena umumnya terdapat lebih dari satu aspek yang dinilai dalam review tersebut. Meskipun penelitian multi-aspek telah dilakukan sebelumnya dan

menunjukkan hasil yang cukup baik, namun masih terdapat ruang untuk pengembangan dengan menggunakan metode *deep learning* dan dataset yang berbeda.

Penelitian ini bertujuan untuk membangun model sentimen multi aspek pada data *review* hotel. Model dibangun dengan memanfaatkan teknologi *deep learning* berbasis LSTM. Hasil pembangunan model diharapkan mampu melakukan analisa informasi lebih mendalam terkait aspek penilaian apa saja yang dibahas dalam sebuah *review*. Sehingga informasi tersebut dapat diolah sebagai bahan untuk pengembangan kualitas hotel guna meningkatkan pelayanan pelanggan.

1.2 Rumusan Masalah

Berdasarkan pembahasan pada latar belakang, maka rumusan masalah penelitian adalah bagaimana membangun model sentimen multi aspek dengan *deep learning* untuk mengetahui penilaian apa saja yang terkandung dalam sebuah *review* hotel?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk membangun model multi aspek pada kalimat *review* hotel. selain itu juga melakukan klasifikasi sentimen pada kalimat *review*. sehingga diharapkan akan dapat memberikan informasi lebih mendalam terkait penilaian yang dibahas dalam sebuah kalimat *review*.

1.4 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Batasan pada penelitian ini adalah dataset yang digunakan adalah domain hotel dengan enam aspek kelas yaitu harga, lokasi, kamar, pelayanan, makanan, dan fasilitas. Masing-masing opini akan dikategorikan dalam sentimen positif, dan negatif.
2. Kombinasi aspek dalam kalimat maksimal tiga. Hal ini didasarkan pada pemahaman bahwa umumnya penulisan kalimat *review* mengandung tidak lebih dari tiga aspek. Pendekatan ini memungkinkan untuk menghindari kompleksitas berlebihan pada proses evaluasi model.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Dapat melakukan identifikasi berbagai aspek yang diulas dalam sebuah kalimat review hotel.
2. Mendapatkan informasi terkait sentimen pada sebuah kalimat review.

1.6 Sistematika Penulisan

Sistematika Penulisan pada penelitian ini akan disusun sebagai berikut:

1. Abstrak berisi rangkuman laporan isi laporan secara umum.
2. Pernyataan Keaslian Tulisan.
3. Daftar Publikasi berisi jurnal internasional.
4. Halaman Kontribusi berisi dosen yang terlibat dalam penelitian.
5. Halaman Persembahan berisi ucapan terhadap keluarga.
6. Kata Pengantar berisi ucapan terimakasih kepada berbagai pihak yang terlibat dalam penelitian.
7. Daftar Isi berisi daftar judul bab, dan sub bab laporan.
8. Daftar Tabel berisi daftar nama tabel.
9. Daftar Gambar berisi daftar nama Gambar.
10. BAB 1 Pendahuluan
 - 1.1 Pendahuluan berisi alasan penulis memilih tema tersebut
 - 1.2 Rumusan Masalah berisi pertanyaan penelitian.
 - 1.3 Tujuan Penelitian berisi tujuan penelitian
 - 1.4 Batasan Masalah berisi batasan apa saja yang akan dibahas dalam penelitian.
 - 1.5 Manfaat Penelitian berisi manfaat yang didapatkan sesudah melakukan penelitian.
 - 1.6 Sistematika Penulisan berisi daftar penulisan dari BAB awal sampai akhir.
11. BAB 2 Tinjauan Pustaka
 - 2.1 Penelitian Terdahulu berisi penelitian yang dilakukan sebelumnya.
 - 2.2 Landasan Teori berisi teori apa saja yang dipakai dalam penelitian.
12. BAB 3 Metodologi
 - 3.1 Data Collection merupakan tahapan pengumpulan dataset review hotel.
 - 3.2 Data Preprocessing adalah Langkah pembersihan dataset sebelum pembuatan model.

3.3 Training Model Data adalah Langkah yang dilakukan untuk membangun model dengan LSTM

3.4 Prediction Evaluation adalah Langkah untuk mengevaluasi model yang dibangun.

13. BAB 4 Hasil Dan Pembahasan

4.1 Proses pembangunan model.

4.2 Proses evaluasi model dengan satu aspek dalam sebuah kalimat.

4.3 Proses evaluasi model dengan kombinasi dua aspek dalam sebuah kalimat.

4.4 Proses evaluasi model dengan kombinasi tiga aspek dalam sebuah kalimat review.

4.5 Pembahasan pada hasil evaluasi terhadap masing-masing proses evaluasi kalimat review.

14. BAB 5 Kesimpulan Dan Saran

5.1 Kesimpulan berisi kesimpulan yang didapatkan sesudah melakukan penelitian.

5.2 Saran berisi saran yang dapat dilakukan pada penelitian selanjutnya.

15. Daftar Pustaka berisi rujukan yang dikutip dalam laporan.

16. Lampiran berisi alamat github yang berisi hasil penelitian yang telah dilakukan.

BAB 2

Tinjauan Pustaka

2.1 Penelitian Terdahulu

Salah satu pendekatan analisis sentimen kalimat diterapkan pada data review film menggunakan model LSTM (*Long Short-Term Memory*) untuk membangun model dan menghasilkan prediksi sentiment review (Widayat, 2021). Dataset yang digunakan sebanyak 25.000 dengan panjang rata-rata per kalimat 233 kata. Dalam penelitian ini metode LSTM dipilih dengan pertimbangan akurasi yang cukup baik saat melakukan pemodelan data teks. Selain itu, LSTM dianggap mampu melakukan pemrosesan data yang relatif panjang. Dengan kombinasi word2vec dan LSTM model yang dibangun dapat mencapai akurasi terbaik sebesar 85.86%. Perolehan hasil akurasi masih bisa dikembangkan dengan mengatur beberapa parameter tuning seperti penambahan jumlah unit LSTM, epoch, maupun batch size.

Penelitian dilakukan (Cahyaningtyas et al., 2021) pada data review hotel dalam Bahasa Indonesia. Proses penelitian dilakukan untuk mengetahui sentimen review dan aspek yang dinilai dalam sebuah kalimat. Beberapa metode *deep learning* diterapkan pada model untuk melihat perbandingan dari sisi akurasi masing-masing metode. Terdapat delapan metode *deep learning* yang digunakan yaitu LSTM, BiLSTM, Attention, GRU, RNN, CNN-LSTM, dan CNN-BiLSTM. Dengan pengaturan arsitektur model berupa 100 unit LSTM, 32 batch size, dan dropout sebesar 0.5 berhasil mendapatkan nilai akurasi tertinggi yaitu sebesar 0.926. Model yang dibangun telah mampu melakukan prediksi terhadap aspek kalimat dan sentimennya. Namun hasil prediksi hanya berlaku untuk kalimat dengan satu aspek saja.

Selanjutnya terdapat pendekatan yang menggunakan model berbasis *Attention* pada sentimen analisis aspek (Liu et al., 2019). Studi ini dilakukan mengingat pentingnya mengetahui emosi maupun alasan seseorang saat mengekspresikan diri melalui sebuah tulisan. Attention-based sentiment reasoner dilakukan pada penelitian ini untuk dapat menganalisa secara presisi terhadap reason atau motif dalam sebuah kalimat review. Domain yang digunakan cukup beragam dari review restaurant, hotel, laptop, kamera, dan lain-lain. Terdapat tiga module dalam pembangunan model yaitu encoder untuk mengambil informasi dari kalimat, reasoner untuk menentukan motif dari kalimat review, dan Aspect-dependent Representation Memory yang berfungsi menyimpan hasil reasoning dan mentransformasi kedalam layer reasoning berikutnya yang memungkinkan.

Penelitian terkait analisis sentimen berbasis aspek pada data teks sudah banyak dilakukan. Secara umum hasil penelitian telah dapat melakukan prediksi terhadap penilaian apa yang terkandung dalam sebuah kalimat. Prediksi dilakukan terhadap kalimat yang hanya mengandung satu aspek atau satu sentiment dalam keseluruhan kalimat. Pada kenyataannya, sebuah kalimat review yang ditulis rata-rata mengandung lebih dari satu penilaian. Sehingga apabila dikembangkan akan mampu memberikan informasi yang lebih banyak. Seperti yang dilakukan (Bouazizi & Ohtsuki, 2017) pada pengenalan aspek dalam kalimat di twitter. Pattern-based approach dilakukan untuk mendapatkan penilaian apa saja yang dibahas dalam sebuah opini. Dengan menggunakan tools SENTA, penelitian ini mengelompokkan opini twitter dengan 7 kelas yaitu *fun*, *love*, *hate*, *neutral*, *angger*, *sadness*, dan *happiness*.

Multi aspek sentimen analisis juga diterapkan pada review terhadap aplikasi tiktok dengan menggunakan metode deep learning (Malik & Sibaroni, 2022). Data review diambil dari google play pada aplikasi Tiktok. Review terhadap aplikasi tiktok dikategorikan kedalam tiga aspek yaitu aspek fitur, konten, dan bisnis. Kombinasi metode deep learning yaitu CNN dikombinasikan dengan TF-IDF diterapkan dalam pembangunan model. Terdapat dua skenario yang dilakukan dalam proses pengujian model yaitu pengujian dengan *preprocessing* dan pengujian dengan *feature expansion*. Hasil pengujian dengan tahapan preprocessing memiliki pengaruh yang tinggi terhadap model CNN. Akurasi setelah dilakukan preprocessing penuh menghasilkan akurasi sebesar 87%.

Penelitian berikutnya terkait multi aspek dilakukan pada sosial media dengan Bahasa Bengali. Penelitian ini dilakukan karena kurangnya performa klasifikasi pada penelitian sebelumnya. Klasifikasi multi aspek dilakukan dengan menggunakan kombinasi metode CNN dan LSTM. Kelas yang dipakai dalam sentiment ini berupa kelas sexual, religious, political, dan acceptable. Dalam pembangunan model, metode CNN dan LSTM digabungkan menjadi CLSTM yang menghasilkan akurasi pada model sebesar 85% dan F1 score sebesar 85%. Model dibangun dengan menggunakan dataset dari facebook sebesar 42,063.

Pada Tabel 2.1 merangkum berbagai penelitian terkait sentimen analisis pada aspek kalimat dengan *Deep Learning*. Tabel terdiri atas sub tema, keywords, ulasan kritis, dan pustaka. Sub tema merupakan bidang penelitian yang dilakukan. Ulasan kritis berisikan intisari dari analisis pengembangan riset yang dilakukan pada penelitian yang menjadi *state of the art*.

Tabel 2.1 Penelitian Terkait *Deep Learning* Pada Analisis Data Teks

Sub Tema	Keywords	Ulasan Kritis	Pustaka
Sentimen analisis movie review	Sentimen analisis, LSTM	Studi terkait sentimen pada review movie dilakukan dengan metode LSTM dan word vector. Uji coba dilakukan dengan mengatur representasi word vector dengan metode CBOW dan Skip-gram. Akurasi diatas 85% diperoleh dari kombinasi LSTM dan word vector.	(Widayat, 2021)
Sentimen analisis berbasis aspek dengan <i>Attention-based</i>	ABSA, <i>Attention-based</i>	Sentimen analisis berbasis aspek dilakukan pada review dengan beragam domain seperti restaurant dan kamera. Metode deep learning yang digunakan adalah Attention-based. Dimana penelitian ini bertujuan untuk mengetahui reason atau motif dan emosi dari penulisan review.	(Liu et al., 2019)
Multi-class sentiment analisis pada opini twitter	<i>Multi-class</i>	Studi terkait pattern-based approach pada opini twitter untuk menentukan aspek multi kelas. Tools SENTA digunakan untuk mengidentifikasi setiap aspek yang terkandung dalam opini twitter	(Bouazizi & Ohtsuki, 2017)
<i>Multi-class sentiment analysis</i>	Multi Aspek, CNN, aplikasi Tiktok	Membandingkan metode <i>deep learning</i> untuk klasifikasi model multi aspek. Terdapat dua skenario pengujian yang dilakukan yaitu pengujian model dengan tahapan preprocessing data, dan tahapan pengujian model tanpa melakukan preprocessing data.	(Malik & Sibaroni, 2022)
<i>Multi-class sentiment analysis</i>	<i>Multi-class</i> , Bengali sosial media, LSTM	Studi dilakukan untuk multi-class sentiment analisis terhadap social media Bahasa Bengali. Pada penelitian ini dilakukan penggabungan metode CNN dan LSTM untuk mendapatkan sentimen multi kelas.	(Haque et al., 2023)
Multi-aspect <i>hate speech detection</i>	<i>Multi-aspect</i> , BERT, <i>hate speech</i>	melakukan <i>training dataset hate speech</i> secara individual dan multi aspek dengan metode BERT untuk deteksi <i>hate speech</i> pada sosial	(Mazari et al., 2023)

		media.	
Sentimen analisis pada hotel review	Sentimen analisis, <i>review</i> hotel, LSTM	Studi terkait sentiment analisis berbasis aspek pada review hotel. Penelitian melakukan perbandingan metode deep learning untuk melakukan prediksi pada aspek dan sentiment review. Metode yang dibandingkan yaitu RNN, CNN, Attention, LSTM, BiLSTM, GRU, CNN-LSTM, CNN-BiLSTM	(Cahyaningtyas et al., 2021)
Sentimen analisis pada hotel review	Sentiment analisis, hotel review, LSTM	Studi dilakukan pada analisis sentiment review hotel berdasarkan aspek yang di review. Tahapan penelitian dimulai dengan mengetahui topic yang dibicarakan dalam review dengan metode LDA. Selanjutnya dilakukan identifikasi atas puas atau tidak puas dengan menggunakan kombinasi metode word embedding dan LSTM.	(Priyantina & Sarno, 2019)
Multi aspek sentimen analisis	Multi aspek, CAN (<i>constrain attention network</i>).	Multi-aspek sentimen analisis dilakukan dengan mengetahui posisi aspek tertentu dalam sebuah kalimat dengan metode <i>constrain attention network</i> (CAN). Secara spesifik ada dua cara yang diterapkan yaitu orthogonal dan sparse untuk melihat posisi multi aspek dalam rangkaian kalimat. Pengujian dilakukan dengan kasus overlapping yaitu identifikasi secara benar multi aspek dengan satu polarity. Dan error case ketika terdapat multiple aspek dengan polarity yang berbeda terhadap masing-masing aspek.	(Hu et al., 2018)

Berdasarkan uraian penelitian terdahulu pada pengolahan data teks dengan menggunakan teknologi *deep learning*, beberapa hipotesis yang berkaitan dengan hasil penelitian adalah :

1. Metode *deep learning* merupakan pendekatan yang efektif untuk mengenali pola yang terkandung dalam setiap kata yang dituliskan dalam kalimat review. Hal ini

memungkinkan metode deep learning untuk mengidentifikasi makna atau motif tersembunyi dalam teks secara lebih mendalam.

2. Secara umum metode deep learning telah banyak digunakan untuk identifikasi single aspek atau sentiment kalimat. Hasil tersebut menjadi acuan untuk melakukan pengembangan identifikasi lebih spesifik terhadap nilai-nilai yang terkandung dalam sebuah kalimat. Mengingat dalam kehidupan nyata, kalimat review dapat mengandung penilaian yang lebih kompleks.
3. Berbagai metode *deep learning* telah dimodelkan dan diuji akurasinya. Beberapa metode yang banyak dicoba seperti CNN, LSTM, GRU, attention, hingga bert. Semua telah dicoba dan dibandingkan hasilnya. Dari tuju penelitian terkait sentiment analisis, empat diantaranya menggunakan metode LSTM untuk proses pengenalan pola pada data teks.

Studi-studi terdahulu telah menunjukkan bahwa penggunaan *deep learning* dalam multi-aspek sentimen analisis dapat menghasilkan kinerja yang unggul dalam mengenali sentimen pada aspek-aspek spesifik. Model *deep learning* ini mampu mengatasi kompleksitas dan variasi dalam teks, dan memberikan hasil yang lebih baik daripada metode tradisional.

2.2 Landasan Teori

2.2.1 Deep Learning

Deep learning adalah sebuah pendekatan dalam pembelajaran mesin (*machine learning*) yang berfokus pada penggunaan arsitektur jaringan saraf tiruan (*neural networks*) yang mendalam untuk memahami dan menganalisis data. Metode ini mencoba untuk mereplikasi proses belajar manusia dengan mengeksplorasi dan memodelkan representasi data yang semakin kompleks dan abstrak.

Metode deep learning mengandalkan jaringan saraf tiruan yang terdiri dari banyak lapisan (layer) yang saling terhubung. Setiap lapisan dalam jaringan ini bertanggung jawab untuk mengambil representasi abstrak dari data yang diberikan dan meneruskan informasi ke lapisan berikutnya, sehingga mampu memahami fitur-fitur yang semakin kompleks dari data tersebut.

Pemodelan menggunakan deep learning telah menunjukkan perkembangan pesat dengan hadirnya berbagai metode yang efektif, seperti Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), dan Long Short-Term Memory (LSTM). Salah

satu metode yang menjadi perhatian khusus adalah LSTM, yang merupakan hasil evolusi dari RNN. RNN memiliki kemampuan untuk memproses data berurutan, tetapi terbatas oleh masalah vanishing gradient, yaitu ketika gradien dari error function menjadi sangat kecil sehingga menyebabkan model sulit beradaptasi dan mempelajari pola pada data dengan baik.

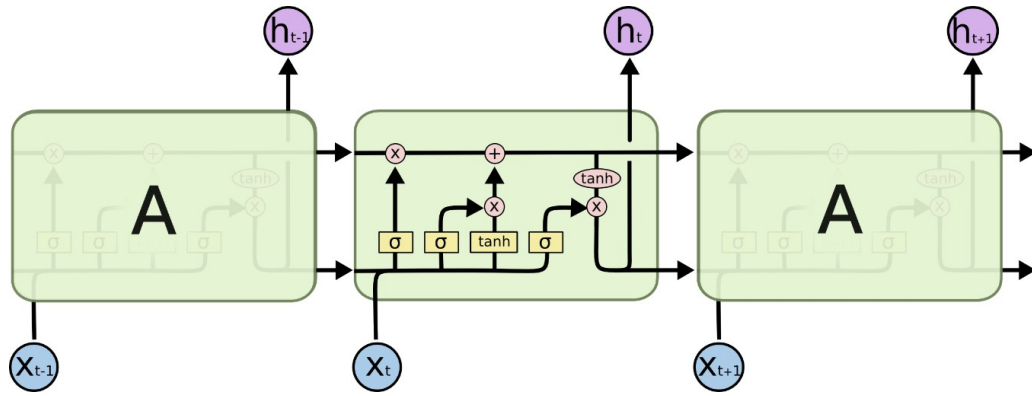
Berbeda dengan RNN, LSTM dirancang khusus untuk mengatasi masalah vanishing gradient ini dengan menggunakan mekanisme pintu (*gate*) yang memungkinkan model untuk mengingat dan mengabaikan informasi sesuai kebutuhan. Oleh karena itu, LSTM menjadi pilihan yang efektif untuk memodelkan data dengan rangkaian panjang, seperti teks, bahasa, dan deret waktu. Implementasi LSTM telah menghasilkan hasil yang menjanjikan dalam berbagai bidang, termasuk pengenalan teks, penerjemahan bahasa, analisis sentimen, dan prediksi deret waktu.

2.2.2 LSTM

Long Short-Term Memory (LSTM) adalah jenis model rekurensi dalam bidang kecerdasan buatan (*artificial intelligence*) yang pertama kali diperkenalkan oleh (Hochreiter & Schmidhuber, n.d.). LSTM dikembangkan untuk mengatasi masalah *vanishing gradient* yang sering terjadi pada model rekurensi tradisional, yang menyebabkan kehilangan informasi yang relevan dalam urutan data yang panjang.

LSTM bekerja dengan menggunakan unit memori khusus yang disebut sel memori (*memory cell*) untuk menyimpan dan mengingat informasi jangka panjang. Untuk lebih jelas, rangkaian memori unit LSTM terlihat pada Gambar 2.1. Sel-sel memori ini dilengkapi dengan gerbang (*gate*) yang mengatur aliran informasi ke dalam dan keluar dari sel memori.

Terdapat tiga jenis gerbang dalam LSTM yaitu gerbang *input* (*input gate*), gerbang lupa (*forget gate*), dan gerbang keluaran (*output gate*) (Zaman et al., 2019). Sistem gerbang ini terdiri dari beberapa arsitektur JST sederhana untuk mengatur perlakuan terhadap setiap data didalamnya. Gerbang-gerbang ini memungkinkan LSTM untuk memilih informasi mana yang akan diingat, digunakan dan dilupakan dalam urutan data.



Gambar 2.1 LSTM arsitektur.

Gambar 2.2 merupakan legenda dari notasi-notasi yang digunakan untuk mengabungkan seluruh proses input hingga output. Lingkaran merah muda mewakili operasi elemen, seperti penambahan atau perkalian elemen vector. Kotak kuning adalah lapis jaringan saraf yang mengandung parameter dan bias. Dua garis yang bergabung menandakan penggabungan dua matriks/vector. Sementara itu, garis berpisah menandakan kontennya disalin dan salinannya pergi ke simpul yang berbeda.



Gambar 2.2 Legenda notasi pada arsitektur LSTM.

Penjelasan lebih lanjut terkait langkah-langkah yang dilalui sebagai *input* dan *output* dari setiap notasi dalam arsitektur LSTM adalah sebagai berikut :

1. X_t adalah input pada langkah waktu (time step) tertentu dalam urutan data. X_t adalah representasi vektor dari token ke- t dalam urutan data dan berperan sebagai input untuk langkah waktu tersebut.
2. h_t adalah output atau representasi dari "hidden state" pada langkah waktu tertentu, yang menyimpan informasi tersembunyi atau terkini yang relevan dari seluruh urutan data sebelumnya hingga langkah waktu ke- t .

3. Tanh adalah fungsi aktivasi yang umum digunakan dalam LSTM untuk memproses input dan menghasilkan output. Fungsi tanh beroperasi pada setiap elemen dari input vektor dan mengubah nilai-nilai tersebut ke dalam rentang -1 hingga 1. Fungsi tanh membantu dalam mengatasi masalah vanishing gradient dan memungkinkan LSTM untuk menyimpan informasi dalam skala yang lebih besar.
4. Cell State (Ct): Cell state atau biasa disebut juga sebagai "memory cell" adalah komponen utama dalam LSTM yang memungkinkan penyimpanan dan pengingatan informasi jangka panjang.

Dengan adanya mekanisme sel memori dan gerbang, LSTM dapat mengatasi masalah *vanishing gradient* dan mengatasi ketergantungan jarak panjang dalam urutan data. Hal ini membuat LSTM sangat efektif dalam memahami konteks dan hubungan jangka panjang dalam data berurutan, seperti teks, suara, atau rangkaian waktu.

Secara keseluruhan, LSTM adalah jenis model rekurensi yang mampu mengingat informasi jangka panjang dan mengatasi masalah vanishing gradient. Kemampuan ini membuat LSTM menjadi sangat berguna dalam banyak aplikasi, termasuk pemrosesan bahasa alami, pemodelan rangkaian waktu, dan analisis sentimen.

2.2.3 Pengaturan *Hyperparameter*

Hyperparameter pada arsitektur model dengan LSTM adalah parameter-parameter yang perlu diatur secara manual sebelum memulai proses training model. Hyperparameter ini tidak dapat dipelajari oleh model selama proses training, melainkan harus ditentukan oleh pengguna berdasarkan eksperimen dan pengetahuan domain. Beberapa contoh hyperparameter pada arsitektur model dengan LSTM antara lain:

1. Jumlah Unit LSTM: Jumlah unit atau neuron di dalam lapisan LSTM. Jumlah ini menentukan kompleksitas model dan dapat mempengaruhi kapasitas model untuk memahami pola dalam data.
2. Jumlah Lapisan LSTM: Model LSTM dapat memiliki beberapa lapisan LSTM yang saling berhubungan secara berurutan. Menambahkan lebih banyak lapisan dapat meningkatkan kompleksitas dan kapasitas model.
3. Jumlah Epochs: Jumlah epoch adalah jumlah iterasi atau putaran penuh dari seluruh data training yang dilakukan selama proses training. Jumlah epoch ini menentukan berapa kali model akan "melihat" data training dan mengupdate bobotnya.

4. Ukuran Batch: Ukuran batch adalah jumlah sampel data yang digunakan dalam satu kali pembaharuan bobot pada model selama proses training. Ukuran batch dapat mempengaruhi kecepatan training dan konvergensi model.
5. Laju Pembelajaran (*Learning Rate*): Laju pembelajaran adalah faktor yang menentukan seberapa besar langkah pembelajaran yang diambil selama proses update bobot pada setiap iterasi. Laju pembelajaran ini dapat mempengaruhi kecepatan konvergensi dan stabilitas model.
6. Dropout Rate: Dropout adalah teknik regulasi yang digunakan untuk mengurangi overfitting. Dropout rate adalah proporsi unit yang akan di-dropout (dinonaktifkan) selama training.
7. Fungsi Aktivasi: Fungsi aktivasi digunakan untuk mengubah keluaran dari unit atau neuron LSTM. LSTM biasanya menggunakan fungsi aktivasi "tanh" untuk pintu forget, input, dan output serta fungsi aktivasi "sigmoid" untuk pintu gate update.
8. Jumlah Step Waktu (Time Steps): Jumlah step waktu adalah jumlah langkah atau urutan data yang digunakan sebagai input untuk model LSTM. Hal ini berhubungan dengan bagaimana data sequential diproses dalam arsitektur LSTM.

Penggunaan hyperparameter yang tepat dapat membantu meningkatkan performa model dan menghindari masalah seperti overfitting atau underfitting. Pada kasus klasifikasi teks pengaturan hyperparameter ini sangat mempengaruhi hasil performa pada model yang dibangun (Veronika Aritonang et al., 2022). Dimana hasil eksperimen dapat terlihat mengalami kenaikan dan penurunan seiring perubahan hyperparameter seperti penambahan jumlah *epochs*, *vector size*, atau *number of word*.

2.2.4 Preprocessing Data

Preprocessing data pada *text analytics* adalah tahap penting dalam pengolahan teks yang dilakukan sebelum analisis teks sebenarnya dimulai. Tujuan dari *preprocessing* adalah memanipulasi data menjadi bentuk yang mudah diolah oleh sistem (Maksun et al., 2021). Beberapa langkah umum dalam *preprocessing* data pada *text analytics* meliputi:

1. Tokenisasi: Memisahkan teks menjadi unit-unit kecil yang disebut "token" atau kata-kata. Tokenisasi memisahkan kalimat menjadi kata-kata terpisah, yang membantu dalam proses analisis dan pemrosesan selanjutnya.

2. *Lowercasing*: Mengubah semua huruf dalam teks menjadi huruf kecil. Hal ini dilakukan untuk menghindari perbedaan kasus (uppercase vs. lowercase) yang dapat mengakibatkan ketidakkonsistenan dalam analisis.
3. Pembersihan (*Cleaning*): Menghapus karakter khusus, tanda baca, angka, atau simbol yang tidak relevan dalam teks. Pembersihan juga melibatkan menghilangkan kata-kata pengisi (stop words) seperti "the," "is," "and," yang sering muncul dalam teks tetapi tidak memberikan informasi penting dalam analisis.
4. Stemming dan Lemmatization: Stemming dan lemmatization adalah teknik untuk mengonversi kata-kata ke bentuk dasar atau kata dasar. Stemming menghapus akhiran kata, sementara lemmatization mengubah kata menjadi bentuk dasar yang lebih bermakna. Contohnya, "running," "runs," dan "ran" dapat diubah menjadi "run."
5. Penghapusan Redundansi: Identifikasi dan penghapusan kata-kata yang mungkin redundan atau memiliki variasi bentuk yang serupa. Misalnya, kata "memasak" dan "memasakkan" dapat dianggap memiliki makna yang sama dan salah satu bentuk dapat dihilangkan.
6. Vektorisasi: Mengubah teks menjadi representasi vektor numerik. Metode umum untuk vektorisasi teks termasuk metode Bag-of-Words (BoW) atau TF-IDF (Term Frequency-Inverse Document Frequency).
7. Encoding: Mengonversi vektor kata menjadi angka, agar dapat diolah oleh model machine learning atau algoritma lainnya.

Preprocessing data pada *text analytics* membantu meningkatkan akurasi dan efisiensi pada model (Malik & Sibaroni, 2022) (Program et al., 2019). Langkah-langkah tersebut juga membantu mengurangi dimensi data dan meningkatkan kualitas representasi teks, sehingga mempermudah penerapan teknik-teknik analisis teks lebih lanjut seperti klasifikasi teks, analisis sentimen, atau penggalian informasi.

2.2.5 Metode pengukuran

Metode pengukuran dilakukan untuk mendapatkan penilaian terhadap hasil pembangunan model. Dalam tahapan ini digunakan dua metode pengukuran yaitu dengan menggunakan confusion matrix dan akurasi kombinasi aspek. Metode pengukuran dengan confusion matrix dilakukan pada pengujian kalimat dengan satu aspek dan sentiment kalimat.

Kemudian metode akurasi kombinasi aspek diterapkan pada penilaian kalimat dengan kombinasi dua dan tiga aspek.

Faktor keberhasilan evaluasi dengan *F1 Score* terjadi ketika label prediksi yang dibandingkan dengan label data aktual memiliki kesesuaian atau kecocokan yang tinggi. *Precision* adalah metrik yang mengukur sejauh mana hasil positif yang diidentifikasi oleh model adalah benar-benar positif. Ini memberikan gambaran tentang ketepatan model dalam mengklasifikasikan positif. *Recall*, di sisi lain, mengukur sejauh mana model dapat menemukan kembali atau mengenali semua kasus positif yang ada dalam data aktual. Perhitungan matriks dilakukan dengan menggunakan (2.1), (2.2), dan (2.3).

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positives} \quad (2.1)$$

$$Recall = \frac{True\ Positive}{True\ Positive, False\ Negative} \quad (2.2)$$

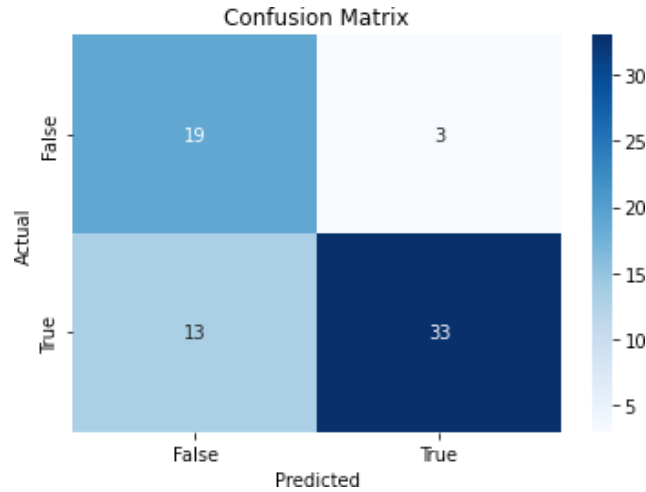
$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (2.3)$$

Sebagai ilustrasi dalam evaluasi model sentimen, penulis mempresentasikan salah satu contoh confusion matrix pada Gambar 2.3. Dalam matriks ini, penulis menggunakan total 68 kalimat ulasan selama pengujian. Hasil pengujian sentimen ini menghasilkan metrik yang menggambarkan performa model, dengan nilai True Positive sebesar 33, True Negative sebanyak 13, False Positive sejumlah 19, dan False Negative hanya 3. Penulis menghitung berbagai metrik evaluasi, seperti yang dijelaskan dalam persamaan (2.1), (2.2), dan (2.3), untuk mendapatkan pemahaman yang lebih baik tentang kinerja model ini. Proses perhitungan adalah sebagai berikut:

$$Precision = \frac{33}{33 + 19} = 0.634$$

$$Recall = \frac{22}{22 + 2} = 0.916$$

$$\begin{aligned}
 F1 &= \frac{2 \times 0.634 \times 0.916}{0.634 + 0.916} \\
 &= 0.749
 \end{aligned}$$



Gambar 2.3 Contoh confusion matrix pada model sentimen.

Evaluasi model pada kalimat dengan kombinasi aspek menggunakan pengukuran akurasi kombinasi K-aspek. Dimana kombinasi akurasi aspek ini bertujuan untuk mengetahui sejauh mana model mampu memprediksikan secara tepat dari kombinasi tiga aspek dan dua aspek secara berturut-turut. Metode ini mengacu pada model perhitungan *Top-K accuracy* sebagai metrik akurasi konvensional dalam klasifikasi.

Top-K accuracy memberikan informasi tentang sejauh mana model mampu melakukan prediksi dengan tepat untuk setiap data. Namun, metrik ini hanya mempertimbangkan prediksi dengan probabilitas tertinggi dan tidak memperhitungkan tingkat ketidakpastian dalam prediksi. Oleh karena itu, untuk beberapa kasus, penggunaan metrik lain seperti top-k accuracy atau log loss dapat memberikan informasi yang lebih komprehensif tentang kinerja model dalam menghadapi ketidakpastian.

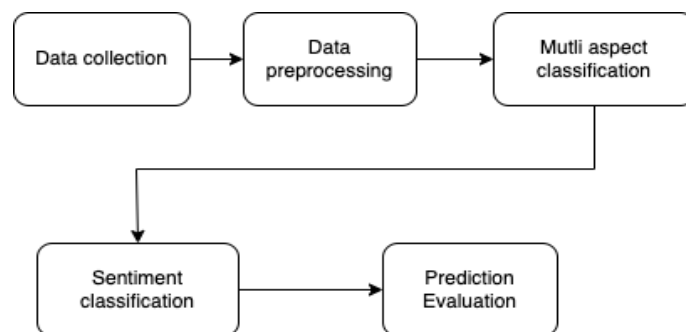
Selanjutnya perhitungan akurasi aspek pada penelitian ini dilakukan menggunakan persamaan (2.4), di mana akurasi diukur berdasarkan jumlah kombinasi (K) aspek dalam suatu kalimat. Jumlah prediksi benar menggambarkan sejauh mana model dapat mengidentifikasi K aspek kombinasi secara tepat. Selanjutnya, total data uji digunakan untuk menentukan jumlah keseluruhan data yang dievaluasi dalam pengujian. Hasil perhitungan ini dinyatakan dalam bentuk persentase untuk memberikan gambaran tentang performa model dalam memprediksi kombinasi aspek.

$$\text{Akurasi } K - \text{ aspek} = \frac{\text{Jumlah prediksi benar}}{\text{jumlah total data uji} * 100} \quad (2.4)$$

BAB 3

Metode Penelitian

Penelitian ini menggunakan metode multi-aspek sentimen analisis untuk menganalisis penilaian pada ulasan hotel. Metodologi yang penulis terapkan melibatkan beberapa tahapan penting, termasuk pengumpulan data ulasan hotel, pembangunan model, dan evaluasi performa model. Dalam pembahasan berikut, penulis akan menjelaskan secara rinci setiap tahap metodologi yang diterapkan. Pada Gambar 3.1 menunjukkan alur penelitian yang akan dilakukan.



Gambar 3.1 Gambaran alur penelitian.

3.1 Data Collection

Dataset yang digunakan dalam penelitian ini diperoleh dari penelitian sebelumnya yang terdiri dari data review hotel (Cahyaningtyas et al., 2021). Dataset terdiri dari tiga elemen utama, yaitu kalimat utama, aspek, dan sentimen. Dalam setiap kalimat utama, terdapat penilaian atau pendapat yang diungkapkan oleh pelanggan terhadap aspek tertentu dari hotel yang ditinjau.

Aspek yang dimaksud merupakan hal-hal spesifik yang menjadi fokus dalam penilaian yaitu pelayanan, lokasi, fasilitas, makanan, harga dan kamar. Selain itu, sentimen juga dicatat untuk menggambarkan apakah ulasan tersebut bersifat positif atau negative terhadap aspek yang dinilai. Contoh dari dataset yang digunakan untuk training model seperti pada Gambar 3.2. Seluruh dataset yang digunakan untuk membangun model dapat dilihat pada file github https://github.com/windiastri/multi-aspect-analysis/tree/main/data_training.

clean text	aspect	sentiment
fasilitas kamar memadai	kamar	negative
interior kamar banyak lampu mati	kamar	negative
kondisi kamar perabotan tua menimbulkan kesan angker	kamar	negative
kamar model tua	kamar	negative
alat_mandi tidak_lengkap tidak_ada sikat_gigi	kamar	negative
toilet ruang nonton tv kotor banyak kotoran rambut	kamar	negative
ada sedikit kotoran sisa tamu sela tempat_tidur	kamar	negative
tirai berlubang kamar	kamar	negative
bathtub bocor toilet air menggenang	kamar	negative
kulkas tidak_berfungsi	kamar	negative
dapat ruangan kotor banyak kecoa	kamar	negative
ac kuno	kamar	negative
air mandi asin	kamar	negative
keamanan baik	lainnya	positive
nyaman betah tidak_ingin pulang	lainnya	positive
tempat nyaman indah berwisata	lainnya	positive
asri	lainnya	positive
anak betul_betul dimanjakan	lainnya	positive
kebersihan diperhatikan	lainnya	positive
bersih	lainnya	positive
kenyamanan keamanan terjaga	lainnya	positive
nyaman seru	lainnya	positive
adem	lainnya	positive

Gambar 3.2 Kumpulan kalimat review dengan aspek utama yaitu kamar.

Selain dataset yang digunakan untuk membangun model, ada juga data evaluasi yang akan digunakan untuk menguji dan mengevaluasi kinerja model yang telah dibuat. Data evaluasi ini diambil langsung dari review hotel yang berasal dari situs pemesanan online. Pendekatan ini memberikan keuntungan dalam menguji model terhadap kalimat review yang secara langsung ditulis oleh pelanggan sesuai dengan pengalaman mereka. Dengan menggunakan data evaluasi yang berasal dari sumber yang aktual dan langsung dari pengguna, model dapat diuji pada skenario yang lebih mirip dengan situasi nyata di dunia nyata.

Data evaluasi yang mencakup review hotel langsung dari situs pemesanan online menjadi bagian penting dalam proses pengujian model. Penggunaan data evaluasi ini akan memberikan pemahaman lebih lanjut tentang kemampuan dan keterbatasan model dalam menganalisis ulasan pelanggan dan menentukan sentimen serta aspek-aspek tertentu dari review hotel. Dengan demikian, penggunaan data evaluasi dapat memberikan wawasan berharga dalam meningkatkan dan mengoptimalkan kinerja model untuk memberikan hasil analisis teks yang lebih akurat dan relevan. Data evaluasi dapat dilihat pada file github https://github.com/windiastri/multi-aspect-analysis/tree/main/data_evaluation.

3.2 Preprocessing Data

Preprocessing data adalah proses krusial dalam pengolahan data yang bertujuan untuk mengolah dan mempersiapkan data agar dapat diinterpretasikan dan diolah dengan baik oleh sistem yang digunakan. Sebagaimana dilakukan oleh (Utami, 2023) dalam penelitian terkait analisis sentiment produk review dari marketplace. *Preprocessing* data membantu untuk menaikkan performa model hingga mencapai akurasi terbaik diangka 90%.

Tahap *preprocessing* ini dilakukan untuk data training dan data validasi secara terpisah. Pada data training, tahapan preprocessing dilakukan untuk mengubah teks menjadi representasi vektor yang dapat dipahami oleh model pembelajaran. Jumlah tahapan *preprocessing* pada data *training* lebih sedikit karena data yang diperoleh telah cukup baik dan relatif bersih, sehingga hanya memerlukan sedikit manipulasi untuk proses pembangunan model.

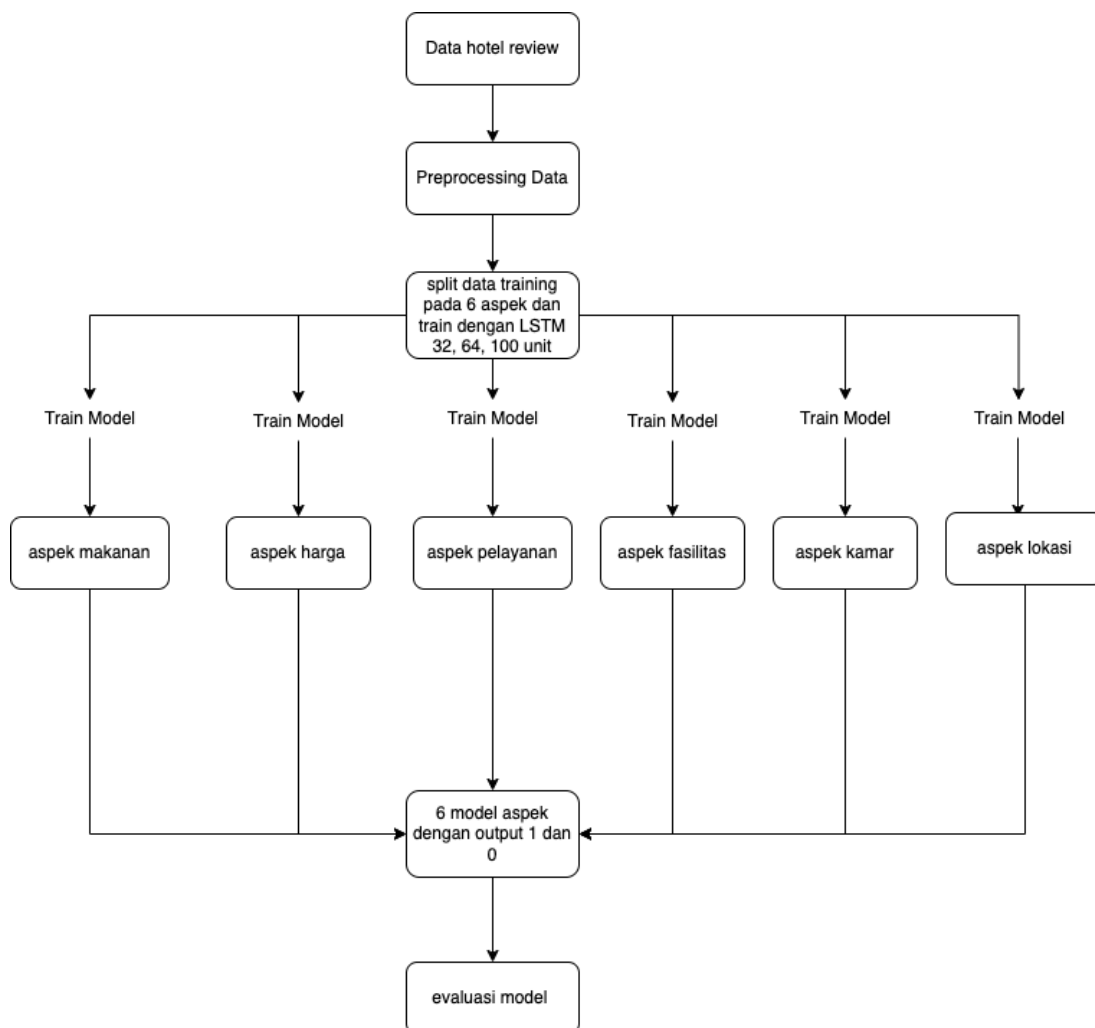
Untuk data uji, dilakukan beberapa tahapan *preprocessing* yang lebih rinci guna mempersiapkan teks secara optimal sebelum dilakukan analisis atau pengujian model. Tahapan *preprocessing* pada data uji mencakup penghapusan tanda baca untuk menghilangkan karakter yang tidak relevan dan mengubah seluruh teks menjadi huruf kecil (*lowercase*) agar kasus tidak mempengaruhi hasil analisis.

	original_text	text_cleaned
1	Kamar yg bersih, fasilitas oke lokasi dekat dg penjual makanan serta pelayanan y...	kamar yg bersih fasilitas oke lokasi dekat dg penjual makanan serta pelayan...
2	Staffnya ramah dan sigap, untuk menu makanan juga termasuk yang lengkap ses...	staffnya ramah dan sigap untuk menu makanan juga termasuk yang lengkap...
3	Super duper ramah all staff.nya. Hotelnya bersih. Recomend	super duper ramah all staff nya hotelnya bersih recommended
4	Sarapan sahurnya juga enak banget rasanya. Lokasi hotel yg bagus banget, di se...	sarapan sahurnya juga enak banget rasanya lokasi hotel yg bagus banget di...
5	Lokasi hotel dekat dengan jalan raya dan ramai, perlu di improve adalah kamar ku...	lokasi hotel dekat dengan jalan raya dan ramai perlu di improve adalah kam...
6	Suka dengan keramahan para staffnya. Kamarnya juga bersih.	suka dengan keramahan para staffnya kamarnya juga bersih
7	Selalu suka nginep disini, kamar luas makanan enak dan variatif bahkan ada indo...	selalu suka nginep disini kamar luas makanan enak dan variatif bahkan ada ...
8	Hotelnya bagus terutama lokasinya sangat oke, banyak sekali makanan sepanjang...	hotelnya bagus terutama lokasinya sangat oke banyak sekali makanan sepa...
9	Pelayanan oke banget, staff nya ramah, parkirannya cukup akomodatif, sarapan bera...	pelayanan oke banget staff nya ramah parkirannya cukup akomodatif sarapan b...
10	Staf nya ramah2 dan sigap, kamar bersih, nyaman, kekurangannya cuma kamar ...	staf nya ramah dan sigap kamar bersih nyaman kekurangannya cuma kam...
11	Kamarnya bersih, wangi, staf nya ramah. Lokasinya strategis dan dikelilingi tem...	kamarnya bersih wangi staf nya ramah lokasinya strategis dan dikelilingi tem...
12	Hotelnya bagus, nyaman, makanan jg enak.	hotelnya bagus nyaman makanan jg enak

Gambar 3.3 Contoh hasil *cleansing* dan *lowercase* pada kalimat review.

Langkah-langkah *preprocessing* penting untuk membersihkan data validasi yang diambil langsung dari review pada situs pemesanan online. Sehingga model dapat menganalisis dan memahami data validasi dengan lebih baik dan akurat. Dengan *preprocessing* yang tepat, diharapkan model dapat memberikan hasil analisis yang lebih relevan pada saat evaluasi.

sentimen. Selain aspek utama, pada setiap file juga terdapat aspek lainnya yang bertujuan untuk membuat model menghasilkan output biner **Error! Reference source not found.** merupakan alur dari proses pembangunan model aspek.

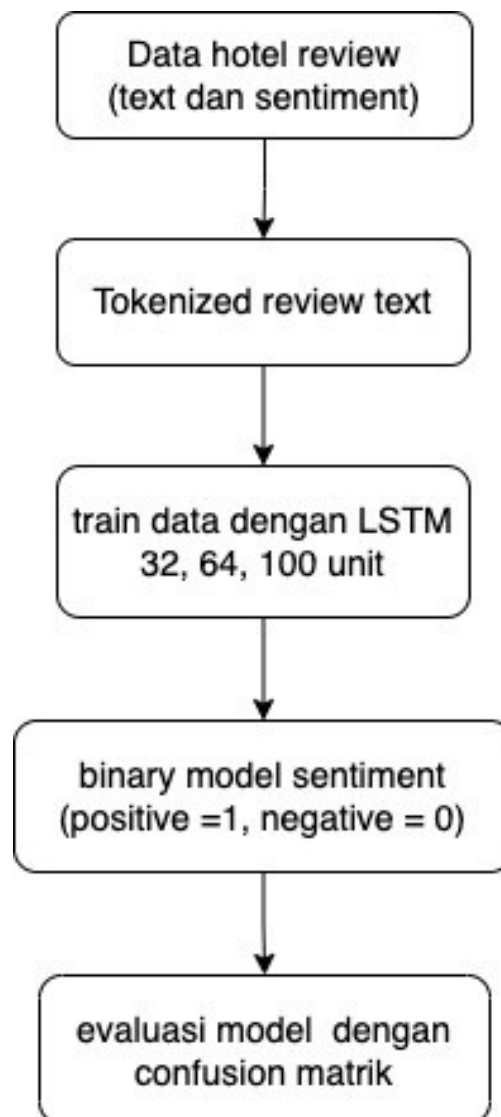


Gambar 3.5 Proses training model aspek.

Pembangunan model untuk setiap aspek melibatkan beberapa langkah. Pertama, menentukan dataset yang kemudian dijadikan sebagai input untuk model LSTM untuk ditraining. Untuk meningkatkan performa model, diterapkan tiga jenis pengaturan unit LSTM yang berbeda yaitu 32, 64, dan 100. Output dari setiap model merupakan nilai binary, di mana angka 1 menunjukkan bahwa kalimat tersebut mengandung aspek yang dimaksud, sedangkan angka 0 menunjukkan ketiadaan aspek tersebut dalam kalimat. Dengan pendekatan ini, penulis dapat mengidentifikasi dan memisahkan setiap aspek dalam kalimat ulasan dengan akurasi yang optimal.

Selanjutnya, Gambar 3.6 adalah tahap pembangunan model sentimen. Proses ini dilakukan dengan menerapkan beberapa langkah yang serupa dengan pengembangan model

aspek sebelumnya. Model sentimen ini dilatih menggunakan seluruh data ulasan dalam satu kali proses, dengan total 5378 data ulasan yang digunakan dalam pelatihan. Hasil output dari model ini berupa penilaian sentimen positif atau negatif. Fokus utama model sentimen adalah untuk melihat penilaian keseluruhan sentimen dalam ulasan tanpa mempertimbangkan aspek-aspek spesifik yang terdapat dalam kalimatnya. Sehingga dalam penelitian ini, sentimen kalimat masih terpisah dari aspek-aspek penilaian dalam rangkaian kalimat review hotel.



Gambar 3.6 Proses training model sentiment.

Dalam upaya untuk mencapai akurasi model yang spesifik dalam memodelkan berbagai aspek dan sentimen kalimat, pembangunan model dilakukan secara terpisah untuk setiap aspek. Pendekatan ini memungkinkan kita untuk lebih mendalam memeriksa dan

memahami sejauh mana metode LSTM mampu mengenali kalimat yang berkaitan dengan masing-masing aspek. Dengan memfokuskan pada masing-masing aspek secara terpisah, kita dapat memperoleh wawasan yang lebih mendalam tentang kemampuan model dalam mengidentifikasi pola dan ciri-ciri unik yang terkait dengan setiap aspek. Dengan begitu, hasilnya diharapkan lebih akurat dan lebih relevan sesuai dengan kebutuhan dan karakteristik setiap aspek yang ingin dimodelkan.

3.3.2 Arsitektur Training Data Review

Selanjutnya, pembuatan arsitektur model tentunya tidak lepas dari hasil penerapan parameter dari beberapa penelitian sebelumnya. Seperti penelitian (Widayat, 2021), tuning parameter pada arsitektur LSTM dibuat dengan susunan LSTM unit 50, epoch 10, batch size 64, dropout 0.5, dan optimizer adam. Hasil akurasi paling rendah dengan susuna paremter tersebut yaitu sebesar 85%.

Pengaturan *hyperparameter* model LSTM juga dilakukan oleh (Cahyaningtyas et al., 2021). Dalam penelitian tersebut susunan parameter yang digunakan yaitu LSTM unit sebesar 100, *batch size* 32, dan *dropout* 0,5. Pengaturan parameter tersebut menghasilkan performa terbaik pada model dengan akurasi sebesar 0.926.

Sehingga pada penelitian ini, penyusunan arsitektur akan dilakukan berdasarkan hasil penelitian sebelumnya dan disesuaikan dengan akurasi saat proses training berlangsung. Dari proses percobaan training data, akan diambil susunan arsitektur yang menghasilkan akurasi terbaik terhadap setiap model. Gambar 3.7 merupakan susunan arsitektur model LSTM yang dibangun untuk model multi aspek sentimen analisis.

```

Model: "sequential_3"
-----
Layer (type)                Output Shape                Param #
-----
embedding_3 (Embedding)     (None, 31, 32)             104992
spatial_dropout1d_3 (Spatia  (None, 31, 32)             0
lDropout1D)
lstm_3 (LSTM)                (None, 64)                 24832
dense_3 (Dense)             (None, 2)                  130
-----
Total params: 129,954
Trainable params: 129,954
Non-trainable params: 0

```

Gambar 3.7 Arsitektur model LSTM.

Terdapat empat layer *sequential* pada model yaitu *embedding*, *SpatialDropout*, LSTM, dan *Dense*. Dari setiap layer memiliki pengaturan parameter yang disesuaikan dengan dataset. Penjelasan lebih lanjut pada masing-masing layer adalah sebagai berikut:

1. Layer *embedding* digunakan untuk mengubah representasi kata ke dalam ruang vektor dengan dimensi yang lebih rendah. Pertama ukuran kamus kata yang digunakan sebesar 3281 berasal dari total kata unik pada dataset. Kedua mengatur dimensi ruang vektor sebesar 32. Ketiga mengatur panjang input teks sebesar 31 yang berasal dari maksimal jumlah kata dalam sebuah kalimat.
2. Menambahkan layer *SpatialDropout1D* untuk mengurangi *overfitting* pada model.
3. Menambahkan layer LSTM untuk memproses urutan data sebesar 64 unit, dengan *dropout rate* 0.4, dan *recurrent dropout rate* 0.2.
4. Menambahkan layer *Dense (fully connected layer)* sebagai layer output dari model sebanyak 2 unit dan fungsi aktivasi sigmoid.

Setelah membuat arsitektur model, langkah selanjutnya adalah menambahkan *optimizer* untuk mengatur proses pelatihan. Sebagaimana dilakukan pada penelitian (Aminullah, n.d.) dan (Widayat, 2021). Dimana kedua penelitian tersebut menggunakan adam optimizer dan binary crossentropy untuk optimasi training data model.

Penggunaan *optimizer* 'adam' dan '*binary_crossentropy*' bertujuan untuk mengurangi *loss function* pada model binary pada pembanguna model multi aspek. Selanjutnya, model akan diterapkan pada dataset *review* hotel dengan melatihnya menggunakan pengaturan 10 *epochs* dan *batch size* sebesar 32. Setiap aspek akan melalui proses pelatihan sehingga terbentuk enam model aspek.

3.3.3 Learning Curve Training Data

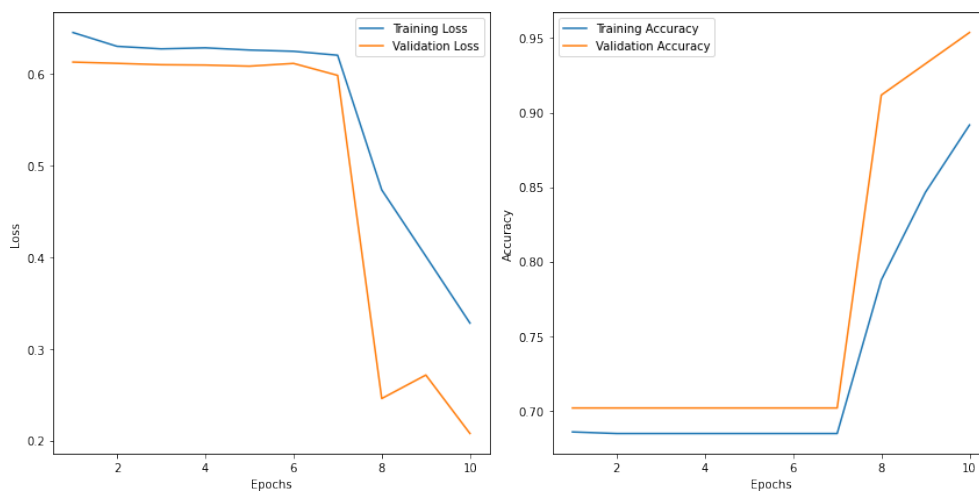
Learning curve menunjukkan bagaimana skor akurasi atau loss dari model berubah seiring dengan jumlah iterasi (*epoch*) pelatihan. Terdapat dua grafik pada setiap proses pembangunan model yaitu grafik loss dan grafik akurasi. Berikut ini penjelasan lebih lanjut terkait Gambar 3.8 yang merupakan salah satu contoh grafik dari aspek makanan:

1. Grafik Loss:
 - Grafik di sebelah kiri menunjukkan perubahan loss (kehilangan) model selama pelatihan.
 - Garis biru mewakili loss pada data pelatihan, sementara garis oranye mewakili loss pada data validasi.

- Jika *Training Loss* dan *Validation Loss* berkurva ke bawah dan semakin mendekati satu sama lain, ini menunjukkan bahwa model pada aspek makanan sedang belajar dan umumnya berfungsi dengan baik.
- Jika *Training Loss* menurun tetapi *Validation Loss* tidak menurun atau bahkan meningkat, ini bisa mengindikasikan *overfitting*, di mana model mempelajari data pelatihan terlalu baik dan tidak generalisasi dengan baik pada data baru (validasi).
- Jika kedua kurva tidak menurun, ini bisa mengindikasikan *underfitting*, di mana model belum cukup kompleks atau kurang dilatih untuk menggambarkan pola dalam data.

2. Grafik Akurasi:

- Grafik di sebelah kanan menunjukkan perubahan akurasi model selama pelatihan.
- Garis biru mewakili akurasi pada data pelatihan, sementara garis oranye mewakili akurasi pada data validasi.
- Jika *Training Accuracy* dan *Validation Accuracy* meningkat dan mendekati satu sama lain, ini menunjukkan bahwa model aspek makanan belajar dengan baik dan umumnya berfungsi dengan baik.
- Jika *Training Accuracy* meningkat tetapi *Validation Accuracy* tidak meningkat atau bahkan menurun, ini bisa mengindikasikan *overfitting*.
- Jika kedua kurva tidak meningkat atau tetap rendah, ini bisa mengindikasikan *underfitting*.



Gambar 3.8 Learning curve model aspek makanan.

Penting untuk mengamati perbedaan antara *Training Loss/Accuracy* dan *Validation Loss/Accuracy*. Jika kurva *Validation Loss/Accuracy* lebih tinggi atau cenderung datar, ini menandakan bahwa model mungkin perlu disempurnakan untuk menghindari *overfitting* atau *underfitting*. Selain itu, perlu juga untuk memperhatikan variasi skor akurasi dan loss dari epoch ke epoch. Jika skor fluktuatif dan tidak stabil, mungkin ada masalah dalam pelatihan atau data yang harus diperhatikan.

Grafik model memberikan gambaran yang cukup baik mengenai proses pembelajaran model. Dengan kedua garis pada grafik loss yang secara konsisten menurun ke bawah, serta pada grafik akurasi yang semakin mendekat satu sama lain menuju puncak. Pola garis yang terbentuk pada grafik tersebut menandakan bahwa tidak terdapat tanda-tanda *underfitting* atau *overfitting* pada saat melatih model aspek makanan.

3.4 Prediction Evaluation

Dalam penelitian ini, penulis menggunakan evaluasi akurasi kombinasi K-aspek pada kombinasi multi aspek dalam kalimat *review* hotel. Tujuan dari penggunaan metode ini adalah untuk mengatasi kemungkinan adanya beberapa aspek yang terdapat dalam satu kalimat *review*. Dalam analisis sentimen berbasis aspek, sebuah kalimat *review* hotel dapat mencakup lebih dari satu aspek, seperti harga, lokasi, pelayanan, makanan, fasilitas, dan kamar. Oleh karena itu, dengan menggunakan metode akurasi kombinasi K-aspek, penulis dapat mengidentifikasi dan memperhitungkan kemungkinan adanya beberapa aspek yang relevan dalam setiap kalimat *review*.

Metode akurasi kombinasi 3 aspek digunakan untuk mengukur sejauh mana model dapat mengidentifikasi tiga aspek utama secara benar dalam kalimat *review*. Model akan memberikan prediksi terhadap aspek yang tepat dalam kalimat tersebut. Sementara itu, metode akurasi kombinasi 2 aspek digunakan untuk melihat sejauh mana model dapat mengenali lebih dua dari tiga yang relevan dalam kalimat *review*. Metode ini mengacu pada perhitungan *Top-K accuracy* dilakukan untuk mengidentifikasi dan memperhitungkan kemungkinan adanya beberapa aspek yang relevan dalam sebuah *review* Click or tap here to enter text.(Simonyan & Zisserman, 2015)Click or tap here to enter text.. Pada Tabel 3.1 merupakan contoh dari proses perhitungan akurasi model.

Tabel 3.1 Contoh Evaluasi Kombinasi K-aspek

Text	Actual aspect	Predicted 3 aspect		Predicted 2 aspect	
Harga tidak sesuai sepadan dengan realita. Kamar terlalu gelap, seperti kamar hotel seharga 150.000an. Sarapan kurang rekomended dan kurang bervariasi.	Harga, kamar, makanan	Harga, kamar, makanan	TRUE	Harga, kamar, lokasi	TRUE
Untuk pelayanan masih standard. Hanya kamarnya tidak ada kualitas, tapi lokasinya lumayan strategis. Dekat dgn stasiun dan jl malioboro.	Pelayanan, kamar, lokasi	Pelayanan, kamar, lokasi	TRUE	Pelayanan, harga, lokasi	TRUE
Tidak sesuai dengan harga hotel. Seharusnya dengan harga tinggi fasilitas yang diberikan dapat lebih. Pelayanan kurang ramah.	Harga, fasilitas, pelayanan	Harga, fasilitas, makanan	FALSE	Harga, fasilitas, lokasi	TRUE
Makananya kurang enak, fasilitas kurang lengkap sabun berdua hanya disediakan 1 shower gel. Tidak sesuai harga.	Makanan, fasilitas, harga	Makanan, lokasi, harga	FALSE	Makanan, fasilitas, pelayanan	TRUE
Menu makanan breakfast terbatas dan sudah banyak yang habis ketika saya masuk slot jam sarapan ke 2 jam 9-10. Overall not recommend dengan harga hotel yang mahal dan pelayanan buruk.	Makanan, harga, pelayanan	Makanan, lokasi, pelayanan	FALSE	Lokasi, fasilitas, pelayanan	FALSE

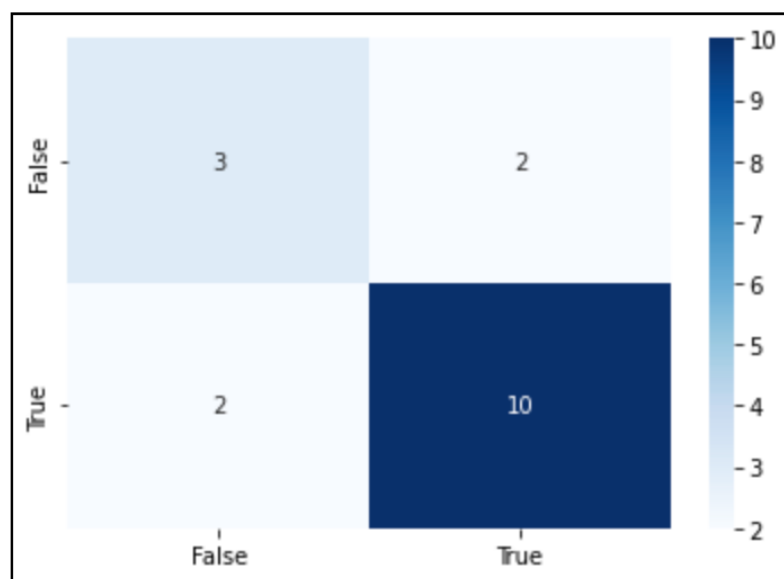
Proses evaluasi model dilakukan secara dua tahap menggunakan persamaan 2.1. Pada tahap pertama, penulis mengukur sejauh mana model mampu mengidentifikasi dan memprediksi tiga aspek yang tepat pada setiap kalimat. Pada tahap kedua, evaluasi dilakukan dengan menghitung jumlah prediksi yang benar ketika model mampu memprediksi kombinasi dua dari tiga aspek yang ada dalam kalimat. Proses perhitungan yang dilakukan dalam kedua tahap ini yaitu :

$$\begin{aligned}
 \text{akurasi 3 – aspek} &= \frac{3}{6 * 100} \\
 &= 50\%
 \end{aligned}$$

$$\begin{aligned}
 \text{akurasi 2 - aspek} &= \frac{5}{6 * 100} \\
 &= 83\%
 \end{aligned}$$

Dalam konteks evaluasi kalimat satu aspek dan sentimen aspek pada kalimat *review* hotel, penulis menggunakan *confussion matrix* untuk mengukur keberhasilan model dalam mengklasifikasikan sentimen berdasarkan aspek yang spesifik. Model dievaluasi dengan membandingkan label prediksi dengan label data actual. Sehingga penulis dapat menilai sejauh mana model dapat memahami dan mengenali sentimen yang sesuai dengan aspek yang dibahas dalam ulasan hotel.

Kemudian penulis menggunakan metode *F1 score* untuk mendapatkan informasi lebih mendalam tentang sejauh mana model mampu menghasilkan prediksi yang akurat dan konsisten dengan data aktual. Nilai *F1 score* yang tinggi menunjukkan kualitas prediksi yang baik. Dimana model mampu mencapai keseimbangan yang baik antara presisi dan *recall*.



Gambar 3.9 Confusion matrix pada analisis sentimen

Dengan menggunakan kedua metode ini, penulis dapat mendapatkan pemahaman yang lebih komprehensif tentang berbagai aspek yang dibahas dalam kalimat *review* hotel. Hal ini memungkinkan penulis untuk mendapatkan informasi yang lebih lengkap dan mendalam mengenai sentimen yang terkait dengan masing-masing aspek dalam ulasan hotel.

BAB 4

Hasil dan Pembahasan

Pada penelitian ini *dataset* terbagi menjadi *dataset* untuk *training* dan validasi seperti pada Tabel 4.1. Data *training* yang digunakan sebanyak 5378 yang merupakan data *review* hotel berbahasa Indonesia. Terdapat kalimat dengan masing-masing 860 aspek harga, 908 aspek lokasi, 864 aspek pelayanan, 962 aspek kamar, 936 aspek fasilitas, dan 856 aspek makanan. Data *training* akan dibagi menjadi dua, 80% sebagai data *training* model dan 20% data testing model.

Tabel 4.1 Jumlah Data Review Pada Masing-Masing Aspek

	harga	lokasi	pelayanan	kamar	fasilitas	makanan
Positive	440	494	440	433	494	437
Negative	420	414	424	529	442	419
5386	860	908	864	962	936	856

Terdapat tiga kategori data validasi untuk pengujian model. Masing-masing kalimat validasi meliputi 60 kalimat satu aspek, 80 kalimat kombinasi dua aspek, dan 20 kalimat kombinasi tiga aspek. Seluruh data validasi diperoleh dari *review* hotel di *platform* pemesanan hotel *online* seperti Tiket.com dan Traveloka.com.

Bagian hasil dan pembahasan ini penulis memaparkan proses lebih rinci untuk mendapatkan hasil pembuatan model dari proses *training*, cara untuk mendapatkan hasil dari matrik konfusi untuk kalimat satu aspek dan akurasi untuk kalimat multi aspek, dan cara untuk mendapatkan hasil matrik konfusi untuk sentimen aspek. Arsitektur model klasifikasi yang digunakan pada penelitian menggunakan metode LSTM.

4.1 Proses Pembangunan Model Multi Aspek

Proses *training* menggunakan model yang dibangun dengan metode LSTM. Training model dilakukan dengan pengaturan ukuran *batch size* 32 dan *Epoch* sebesar 10. Gambar 4.1 memuat *Library* yang dibutuhkan dan terdapat penjelasan *library* secara lebih terperinci.

```

import pandas as pd
import numpy as np
# Libraries for Text Preprocessing
import re
from keras.preprocessing.text import Tokenizer
from keras.preprocessing.sequence import pad_sequences
# Libraries for Model
from keras.models import Sequential
from keras.layers import Flatten, Dense, Embedding, LSTM, SpatialDropout1D
from sklearn.model_selection import train_test_split
from sklearn import metrics
import seaborn as sns

tokenizer=Tokenizer()

```

Gambar 4.1 *Import library training data.*

- a. *Import pandas* : *library pandas* manipulasi dan analisis data tabular.
- b. *Import numpy* : digunakan untuk operasi numerik dan manipulasi *array*.
- c. *Import re*: digunakan untuk melakukan manipulasi dan pencocokan pola pada string dengan menggunakan *regular expression*.
- d. *Import Tokenizer*: digunakan untuk membagi teks menjadi token-token yang dapat digunakan dalam pemrosesan teks selanjutnya.
- e. *Import pad_sequences*: digunakan untuk mengisi atau memotong urutan (misalnya daftar token) menjadi panjang yang sama.
- f. *Import Sequential*: digunakan untuk membuat model *neural network* secara sekuensial.
- g. *Import Flatten, Dense, Embedding, LSTM, SpatialDropout1D*: Mengimpor beberapa *class layer* untuk membangun arsitektur model neural network.
- h. *Import train_test_split*: Mengimpor fungsi *train_test_split* untuk membagi dataset menjadi subset train dan test.
- i. *From sklearn import metrics*: modul ini menyediakan berbagai metrik evaluasi untuk mengukur performa model.
- j. *Import seaborn*: digunakan untuk visualisasi data yang lebih menarik dan informatif.

Gambar 4.2 merupakan proses *import dataset* yang akan digunakan sebagai data training pada model masing-masing aspek dan model sentimen. Dataset disimpan kedalam file excel. Untuk melakukan pengecekan pada data dalam file digunakan *code data_set.head()*.


```
data_set = pd.read_excel('aspek pelayanan.xlsx')
data_set['text'] = data_set['clean text']
data_set.head()
```

Gambar 4.2 *Import data training set.*

Gambar 4.3 adalah proses tokenisasi dan pengubahan teks menjadi urutan angka yang dapat digunakan sebagai input untuk model.

- a. `tokenizer.fit_on_texts(data_set["text"].values)`: Membuat objek *tokenizer* menggunakan *library* Keras dan melatihnya pada teks yang terdapat dalam kolom 'text' dari dataframe `data_set`. Ini akan menghasilkan indeks kata unik dari teks yang digunakan untuk membangun kamus kata-kata yang akan digunakan selama proses tokenisasi.
- b. `X_fasilitas = tokenizer.texts_to_sequences(data_set["text"].values)`: Mengonversi teks pada kolom 'text' dari dataframe `data_set` menjadi urutan angka dengan menggunakan tokenizer yang sudah dilatih sebelumnya. Setiap kata dalam teks akan digantikan dengan indeks numerik yang sesuai.
- c. `X_fasilitas = pad_sequences(X_fasilitas, maxlen=31, padding='post')`: Melakukan proses padding atau pengisian nol pada urutan angka yang dihasilkan sebelumnya, sehingga semua urutan memiliki panjang yang sama. Parameter `maxlen=31` menentukan panjang maksimum urutan, dan `padding='post'` menunjukkan bahwa nol akan ditambahkan setelah urutan angka.
- d. `vocab_size = len(tokenizer.word_index) + 1`: Menghitung ukuran kosakata atau jumlah kata unik dalam kamus yang telah dibangun dengan tokenizer. `tokenizer.word_index` mengembalikan kamus dengan kata-kata dan indeks numeriknya. Ditambahkan 1 untuk mengakomodasi indeks 0 yang digunakan untuk padding.

```
tokenizer.fit_on_texts(data_set["text"].values)
X_fasilitas=tokenizer.texts_to_sequences(data_set["text"].values)
X_fasilitas=pad_sequences(X_fasilitas, maxlen=31, padding='post')
vocab_size = len(tokenizer.word_index) + 1
```

Gambar 4.3 Melakukan tokenisasi pada data teks.

Gambar 4.4 menunjukkan proses inisiasi aspek yang akan digunakan sebagai label prediksi. Pada tahap ini dilakukan standar label dengan urutan pada kolom pertama label

yang dituju dan kolom kedua sebagai label lainnya. Setelah mendapatkan hasil label aspek, kemudian dilakukan pemisahan data *training* dan testing pada masing-masing *dataset*.

```
y=pd.get_dummies(data_set[["aspect"]])

# switch values for data aspect makanan
new_column_order = ['aspect_pelayanan', 'aspect_lainnya']
y = y[new_column_order]

# Split the data into training and testing sets
x_train,x_test,y_train,y_test = train_test_split(X_fasilitas, y, test_size=0.2, random_state=42)
print("X Train: ",X_kamar)
print("Y Train: ",y_train.shape)
print("X Test: ",x_test.shape)
print("Y Test: ",y_test.shape)
print(y)
```

Gambar 4.4 Mengambil aspek pada *training* data.

Selanjutnya Gambar 4.5 merupakan tahapan terkait dengan pembuatan, konfigurasi, dan pelatihan model LSTM untuk analisis sentimen aspek. Pembuatan model pada masing-masing aspek dan sentimen dibuat terpisah. Sehingga terdapat total 7 model yang terdiri dari enam model aspek dan satu model sentimen. Berikut adalah penjelasan masing-masing baris kode:

- a. *Import tensorflow as tf* : memanggil *library TensorFlow* yang digunakan untuk membangun dan melatih model *deep learning*.
- b. *model = Sequential ()* : Membuat objek model *sequential*, yang akan digunakan untuk menambahkan lapisan-lapisan secara berurutan.
- c. *model.add(Embedding(3281, 32, input_length=31))* : Menambahkan lapisan *embedding* sebagai lapisan pertama dalam model. *Embedding* digunakan untuk mengkonversi urutan angka menjadi representasi vektor yang lebih padat. Parameter `3281` menunjukkan ukuran kamus (jumlah kata unik), `32` adalah dimensi *embedding* yang dihasilkan, dan `input_length=31` menentukan panjang urutan *input*.
- d. *model.add(SpatialDropout1D(0.8))* : Menambahkan lapisan *Spatial Dropout* untuk mengurangi *overfitting*. *Dropout* secara acak mengnonaktifkan sebagian unit dalam lapisan sebelumnya. Parameter `0.8` menunjukkan tingkat *dropout*, yaitu 80% unit akan dinonaktifkan selama pelatihan.
- e. *model.add(LSTM(64, dropout=0.4, recurrent_dropout=0.2))* : Menambahkan lapisan LSTM sebagai lapisan berikutnya dalam model. Parameter `64` menunjukkan jumlah unit LSTM dalam lapisan ini, `dropout=0.4` menunjukkan

tingkat *dropout* untuk unit LSTM, dan ``recurrent_dropout=0.2`` menunjukkan tingkat *dropout* untuk koneksi rekursif dalam LSTM.

- f. `model.add(Dense(2, activation='sigmoid'))` : Menambahkan lapisan *Dense* sebagai lapisan *output* terakhir dalam model. Lapisan ini memiliki 2 unit yang menghasilkan *output* dengan fungsi aktivasi *sigmoid*, yang cocok untuk masalah klasifikasi biner.
- g. `model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['acc'])` : Mengompilasi model dengan menggunakan optimizer 'adam', *loss function* 'binary_crossentropy' (untuk masalah klasifikasi biner), dan metrik evaluasi akurasi ('acc').
- h. `history = model.fit(x_train, y_train, epochs=10, batch_size=32, validation_data=(x_test, y_test))` : Melakukan pelatihan model menggunakan data pelatihan (``x_train`` dan ``y_train``). Pelatihan dilakukan dengan 10 *epoch* dan ukuran *batch* sebesar 32. Data validasi digunakan untuk melakukan evaluasi model pada setiap *epoch* untuk memantau performa model selama pelatihan.

```
import tensorflow as tf

# vocab_size = len(tokenizer.word_index) + 1

model = Sequential()
model.add(Embedding(3281, 32, input_length = 31))
model.add(SpatialDropout1D(0.8))
model.add(LSTM(64, dropout=0.4, recurrent_dropout=0.2))
model.add(Dense(2, activation='sigmoid'))
model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['acc'])
history = model_pelayanan.fit(
    x_train,
    y_train,
    epochs=10,
    batch_size=32,
    validation_data=(x_test, y_test)
)
model.summary()
```

Gambar 4.5 Melakukan proses pemodelan data.

Gambar 4.6 adalah proses untuk melakukan visualisasi learning curve dari training model. Pada proses ini terdapat dua grafik yaitu grafik loss dan grafik akurasi. Berikut adalah penjelasan terkait proses yang dilakukan :

1. `train_loss = history.history['loss']`: Mengambil nilai loss pada setiap epoch dari riwayat pelatihan model dan menyimpannya dalam variabel `train_loss`.
2. `val_loss = history.history['val_loss']`: Mengambil nilai loss pada setiap *epoch* dari riwayat validasi model dan menyimpannya dalam variabel `val_loss`.
3. `train_acc = history.history['accuracy']`: Mengambil nilai akurasi pada setiap epoch dari riwayat pelatihan model dan menyimpannya dalam variabel `train_acc`.

4. `val_acc = history.history['val_accuracy']`: Mengambil nilai akurasi pada setiap *epoch* dari riwayat validasi model dan menyimpannya dalam variabel `val_acc`.
5. `epochs = range(1, num_epochs+1)`: Membuat rentang nilai untuk sumbu x pada grafik, dimulai dari *epoch* pertama hingga *epoch* terakhir.
6. `plt.figure(figsize=(12, 6))`: Membuat sebuah figure (gambar) dengan ukuran 12x6 inci sebagai wadah untuk kedua grafik yang akan ditampilkan.
7. `plt.subplot(1, 2, 1)`: Membuat subplot pertama dengan konfigurasi 1 baris dan 2 kolom, dan memilih subplot pertama sebagai tempat untuk menampilkan grafik *loss*.
8. `plt.plot(epochs, train_loss, label='Training Loss')`: Menampilkan garis untuk grafik *loss* pada data pelatihan dengan menggunakan nilai *epoch* sebagai sumbu x dan *train_loss* sebagai nilai *loss* pada sumbu y. Memberikan label *Training Loss* untuk legenda.
9. `plt.plot(epochs, val_loss, label='Validation Loss')`: Jika data validasi tersedia, maka menampilkan garis untuk grafik *loss* pada data validasi. Nilai *epoch* sebagai sumbu x dan *val_loss* sebagai nilai *loss* pada sumbu y. Memberikan label *Validation Loss* untuk legenda.
10. `plt.xlabel('Epochs')`: Memberikan label pada sumbu x dengan teks "Epochs".
11. `plt.ylabel('Loss')`: Memberikan label pada sumbu y dengan teks "Loss".
12. `plt.legend()`: Menampilkan legenda yang berisi label "Training Loss" dan "Validation Loss" untuk membedakan kedua garis pada grafik.
13. `plt.subplot(1, 2, 2)`: Membuat subplot kedua dengan konfigurasi 1 baris dan 2 kolom, dan memilih subplot kedua sebagai tempat untuk menampilkan grafik akurasi.
14. `plt.plot(epochs, train_acc, label='Training Accuracy')`: Menampilkan garis untuk grafik akurasi pada data pelatihan dengan menggunakan nilai *epoch* sebagai sumbu x dan *train_acc* sebagai nilai akurasi pada sumbu y. Memberikan label "Training Accuracy" untuk legenda.
15. `plt.plot(epochs, val_acc, label='Validation Accuracy')`: Jika data validasi tersedia, maka menampilkan garis untuk grafik akurasi pada data validasi. Nilai *epoch* sebagai sumbu x dan *val_acc* sebagai nilai akurasi pada sumbu y. Memberikan label "Validation Accuracy" untuk legenda.
16. `plt.xlabel('Epochs')`: Memberikan label pada sumbu x dengan teks "Epochs".
17. `plt.ylabel('Accuracy')`: Memberikan label pada sumbu y dengan teks "Accuracy".
18. `plt.legend()`: Menampilkan legenda yang berisi label "Training Accuracy" dan "Validation Accuracy" untuk membedakan kedua garis pada grafik.

19. `plt.tight_layout()`: Mengatur tata letak plot agar tidak tumpang tindih.
20. `plt.show()`: Menampilkan kedua grafik learning curve dalam satu tampilan visual.

```
● # create learning curve each model
train_loss = history.history['loss']
val_loss = history.history['val_loss']
train_acc = history.history['acc']
val_acc = history.history['val_acc']

epochs = range(1, 10+1)

plt.figure(figsize=(12, 6))
plt.subplot(1, 2, 1)
plt.plot(epochs, train_loss, label='Training Loss')
plt.plot(epochs, val_loss, label='Validation Loss')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.legend()

plt.subplot(1, 2, 2)
plt.plot(epochs, train_acc, label='Training Accuracy')
plt.plot(epochs, val_acc, label='Validation Accuracy')
plt.xlabel('Epochs')
plt.ylabel('Accuracy')
plt.legend()

plt.tight_layout()
plt.show()

✓ 0.2s
```

Gambar 4.6 Proses visualisasi learning curve training model.

Gambar 4.7 adalah proses menyimpan model data yang telah di *training* dengan `model.save()`. Sebelum model disimpan, dilakukan pengujian secara sederhana untuk memastikan model yang dibangun telah sesuai ekspektasi. Proses pengujian ini menggunakan satu kalimat saja.

```
review_test = ['staff melayani dengan ramah']
review_tok=tokenizer.texts_to_sequences(review_test)
review_pad=pad_sequences(review_tok,maxlen=31,padding='post')
prediction=model_pelayanan.predict(review_pad).flatten()
model_pelayanan.save('aspek_pelayanan_3.h5')
```

Gambar 4.7 Menyimpan model data yang telah dilakukan.

4.2 Proses Evaluasi Model dengan Kalimat Satu Aspek

Setelah melakukan pemodelan data pada seluruh data aspek dan dengan arsitektur LSTM, data model disimpan dalam format .h5. Kemudian setiap model aspek akan dilakukan evaluasi dengan kombinasi kalimat satu aspek, kombinasi kalimat dua aspek, dan kalimat kombinasi tiga aspek. Berikut adalah penjelasan lebih detail terkait kode evaluasi model:

- a. *Import pandas as pd*: untuk manipulasi dan analisis data.
- b. *Import numpy as np*: untuk operasi numerik dan array multidimensi.
- c. *Import re*: untuk pemrosesan teks, khususnya melakukan pembersihan teks.
- d. *Import tensorflow as tf*: library TensorFlow untuk membangun dan melatih model *deep learning*.
- e. *Import Tokenizer*: melakukan konversi teks data validasi menjadi urutan angka (token).
- f. *Import pad_sequences*: untuk melakukan padding atau *truncation* pada urutan angka pada kalimat review data validasi.
- g. *From sklearn import metrics*: untuk mengukur metrik evaluasi F1_Score.

```
import pandas as pd
import numpy as np
# Libraries for Text Preprocessing
import re
import tensorflow as tf
from keras.preprocessing.text import Tokenizer
from keras.preprocessing.sequence import pad_sequences
from sklearn import metrics

tokenizer=Tokenizer()
```

Gambar 4.8 *Import library* untuk proses validasi.

Tahapan selanjutnya yaitu mempersiapkan data evaluasi pada klasifikasi satu aspek. Pada evaluasi ini digunakan total 70 data uji untuk kalimat dengan satu aspek. Setelah data evaluasi dipersiapkan, kemudian dilakukan tokenisasi data evaluasi. hal ini dimaksudkan agar data teks dapat dibaca oleh model dalam bentuk angka.

Gambar 4.9 menunjukkan *code* proses inisiasi data evaluasi kalimat satu aspek. Pada tahap inisiasi data kalimat satu aspek, dilakukan *cleansing* teks. Hal ini bertujuan untuk menyamakan struktur data teks dengan model yang dibangun. Setiap kalimat yang masuk akan diberi proses hapus special karakter, menghapus angka, spasi, dan *lowercase*.

```

import neatttext.functions as nfx

data_validation = pd.read_excel('single_mini (2).xlsx')
data_validation.tail()

def cleaning(text):
    text = re.sub(r'\n', ' ', text) # Hapus \n (enter)
    text = nfx.remove_hashtags(text) # Hapus hashtags
    text = nfx.remove_numbers(text) # Hapus number
    text = text.strip() # Hapus Whitespace
    text = re.sub('[^0-9a-zA-Z]+', ' ', text) # Hapus karakter selain alfabet dan angka
    return text

def casefolding(text):
    return text.lower()

data_validation['content_cleaned'] = data_validation['text'].apply(cleaning).apply(casefolding)

data_validation.head()

```

Gambar 4.9 Inisiasi data evaluasi.

Selanjutnya dilakukan proses tokenisasi data teks pada baris *code* Gambar 4.10. tokenisasi dilakukan dengan mengatur panjang maksimum input kata sebesar 31 kata. Kemudian ditambahkan padding ‘post’ untuk mengisi kalimat dengan angka 0 apabila kalimat kurang dari ukuran 31 kata.

```

tokenizer.fit_on_texts(data_validation['content_cleaned'])
X=tokenizer.texts_to_sequences(data_validation['content_cleaned'])
X=pad_sequences(X, maxlen=31, padding='post')

```

Gambar 4.10 Melakukan tokenisasi pada data evaluasi kalimat satu aspek.

Proses Gambar 4.11 adalah validasi model dengan data evaluasi yang telah dipersiapkan. Proses validasi menggunakan kalimat satu aspek. Model akan melakukan prediksi terhadap satu aspek pada data evaluasi tersebut. Hasil pengujian model akan disimpan kedalam *variabel prediction_aspect_result*.

```

from keras.models import load_model

threshold = 0.5
prediction_result = []
model = load_model('aspek_fasilitas.h5')

predictions_label = model.predict(X)
predictions_label = tf.where(predictions_label < threshold, 0, 1)

for result in predictions_label:
    prediction_result.append(result[0].numpy())

print(prediction_result)

```

Gambar 4.11 Melakukan evaluasi pada model dengan data teks.

Hasil prediksi kalimat satu aspek yang diperoleh akan dihitung nilai akurasi dengan menggunakan *confusion matrix*. Pengujian ini akan dilakukan sebanyak enam kali sesuai jumlah aspek yang digunakan dalam penelitian ini. Sehingga terdapat enam matriks dan hasil

F1 score untuk seluruh aspek. Gambar 4.12 dan Gambar 4.13 merupakan proses untuk memperoleh nilai dari F1 score.

```
from sklearn.metrics import confusion_matrix
import seaborn as sns
import matplotlib.pyplot as plt

data = data_validation['makanan'].tolist()
actual_aspect = [int(x) for x in data]

# Membuat confusion matrix
confusion_mat = confusion_matrix(prediction_result, actual_aspect)

confusion_df = pd.DataFrame(confusion_mat, index=['False', 'True'], columns=['False', 'True'])
print(confusion_df)

# Visualisasi confusion matrix
sns.heatmap(confusion_df, annot=True, cmap='Blues')
plt.xlabel('Predicted')
plt.ylabel('Actual')
plt.title('Confusion Matrix')
plt.show()
```

Gambar 4.12 Melakukan konfusi matriks pada pengujian kalimat satu aspek.

```
y_pred = np.array(prediction_result)
y_true = np.array(actual_aspect)

recall = metrics.recall_score(y_true, y_pred)
precision = metrics.precision_score(y_true, y_pred)
accuracy = metrics.accuracy_score(y_true, y_pred)
F1_score = metrics.f1_score(y_true, y_pred)
print("recall" ,recall)
print("precision" ,precision)
print("accuracy" ,accuracy)
print("F1_score" ,F1_score)
```

✓ 0.0s

Gambar 4.13 Melakukan perhitungan F1_Score pada hasil prediksi.

4.3 Proses Evaluasi Model dengan Kalimat Multi Aspek

Proses evaluasi multi aspek diterapkan langkah yang sama seperti tahapan evaluasi pada kalimat satu aspek. Hanya saja pada multi aspek terdapat perbedaan jumlah model pengujian dan perhitungan akurasi. Selain itu, pada pengujian ini digunakan metode penilaian akurasi prediksi akurasi kombinasi K-aspek.

Tujuan dari akurasi kombinasi K-aspek pada perhitungan akurasi hasil prediksi multi aspek adalah untuk mengevaluasi sejauh mana model berhasil dalam memprediksi aspek yang tepat dari suatu kalimat atau teks. Penilaian akurasi ini mengacu pada proporsi

persentase prediksi yang benar di mana aspek dengan nilai prediksi tertinggi adalah aspek yang sebenarnya.

Dalam konteks multi-aspek sentimen analisis, ada beberapa aspek yang harus diprediksi dari setiap kalimat atau teks. Misalnya, jika ada enam aspek yang mungkin berupa harga, lokasi, pelayanan, makanan, fasilitas, dan kamar, akurasi kombinasi aspek-3 akan mengukur seberapa sering model berhasil memprediksi aspek yang benar sebagai aspek dengan nilai prediksi tertinggi.

Dengan menggunakan akurasi K-aspek, kita dapat menilai kinerja model dalam mengenali dan memprediksi aspek yang tepat dari suatu teks. Semakin tinggi nilai akurasi K-aspek, semakin baik model dalam mengidentifikasi aspek yang relevan dari teks tersebut. Gambar 4.14 menunjukkan tahapan evaluasi data dengan kombinasi dua aspek. Terdapat 30 kelas kombinasi dua aspek yang berbeda. Pada Gambar 4.14 merupakan kombinasi untuk aspek pelayanan dan aspek lokasi.

```
model_filenames = ['aspek_pelayanan.h5', 'aspek_lokasi.h5'] # List of models
models = []
for filename in model_filenames:
    model = load_model(filename)
    models.append(model)

# Lakukan prediksi pada seluruh data test menggunakan setiap model
predictions = np.zeros((X_validasi.shape[0], len(models)))

for i, model in enumerate(models):
    preds = model.predict(X_validasi)
    # print(preds)
    preds_binary = np.where(preds > 0.5, 1, 0)
    for j, pred in enumerate(preds_binary):
        predictions[j][i] = pred[0]

prediction_aspect_result = predictions.astype(int)

# predictions.dtype
# print(prediction_aspect_result)
# print(data_set['text'])
```

Gambar 4.14 Melakukan evaluasi dua aspek.

Setelah tahapan evaluasi kombinasi dua aspek, diperoleh hasil prediksi yang kemudian akan dihitung akurasi. Data evaluasi telah dilakukan labeling secara manual dengan ketentuan yaitu aspek makanan = 1, aspek fasilitas = 2, aspek pelayanan = 3, aspek kamar = 4, aspek harga = 5, dan aspek lokasi = 6. Pada tahapan pengujian akurasi, masing-masing label aspek akan diubah kedalam bentuk 0 dan 1. angka 0 menunjukkan tidak adanya aspek dalam kalimat, sedangkan angka 1 menunjukkan bahwa terdapat salah satu dari 6

aspek dalam suatu kalimat. Rangkaian proses evaluasi dapat terlihat pada baris *code* Gambar 4.15.

```
#makanan=1,fasilitas=2,pelayanan=3,kamar=4,harga=5,lokasi=6
new_df = data_set[['pelayanan', 'lokasi']]
actual_aspect = new_df.values.tolist()
actual_aspect = data = [[1 if val == 3.0 or val == 6.0 else 0 for val in row] for row in actual_aspect]

y_pred = np.array(prediction_aspect_result)
y_true = np.array(actual_aspect)

# Menghitung Top-1 Accuracy
y_pred_top1 = np.argmax(y_pred, axis=1)
top1_accuracy = np.mean(np.equal(y_pred_top1, np.argmax(y_true, axis=1)))

print("Top-1 Accuracy:", round(top1_accuracy, 1))
Top-1 Accuracy: 0.7
```

Gambar 4.15 Melakukan perhitungan akurasi pada dua aspek.

Selanjutnya Gambar 4.16 adalah proses evaluasi model dengan kalimat kombinasi tiga aspek. Pada evaluasi kombinasi tiga aspek, menerapkan tahapan yang sama dengan evaluasi dua aspek. Hanya saja terdapat perbedaan pada jumlah model aspek yang digunakan dalam satu kali pengujian.

```
from keras.models import load_model

model_filenames = ['aspek_kamar.h5', 'aspek_makanan.h5', 'aspek_pelayanan.h5'] # List of models
models = []
for filename in model_filenames:
    model = load_model(filename)
    models.append(model)

# Lakukan prediksi pada seluruh data test menggunakan setiap model
predictions = np.zeros((X_validasi.shape[0], len(models)))

for i, model in enumerate(models):
    preds = model.predict(X_validasi)
    preds_binary = np.where(preds > 0.5, 1, 0)
    for j, pred in enumerate(preds_binary):
        predictions[j][i] = pred[0]

prediction_aspect_result = predictions.astype(int)

print(prediction_aspect_result)
```

Gambar 4.16 Melakukan evaluasi kalimat kombinasi tiga aspek.

Perhitungan akurasi pada kalimat kombinasi tiga aspek menggunakan akurasi kombinasi K-aspek. hal ini dilakukan untuk bisa melihat kemungkinan prediksi pada aspek utama dan beberapa aspek pendukung dalam kalimat kombinasi 3 aspek. Akurasi 3-aspek memberikan gambaran tentang keberhasilan prediksi tiga aspek utama secara tepat, sedangkan akurasi 2-aspek memberikan informasi tentang kemampuan model dalam

mengenali beberapa dua aspek yang relevan dari teks, meskipun aspek utama tidak dikenali dengan benar. Gambar 4.17 adalah *code* tahapan perhitungan akurasi model untuk kalimat kombinasi tiga aspek.

```
new_df = data_set[['kamar', 'makanan', 'pelayanan']]
actual_aspect = new_df.values.tolist()
actual_aspect = data = [[1 if val != 0.0 else 0 for val in row] for row in actual_aspect]

print(actual_aspect)

y_pred = np.array(prediction_aspect_result)
y_true = np.array(actual_aspect)

# Menghitung Top-1 Accuracy
y_pred_top1 = np.argmax(y_pred, axis=1)
top1_accuracy = np.mean(np.equal(y_pred_top1, np.argmax(y_true, axis=1)))

# Menghitung Top-2 Accuracy
y_pred_top2 = np.argsort(y_pred, axis=1)[: , -2:]
top2_accuracy = np.mean(np.any(y_pred_top2 == np.argmax(y_true, axis=1)[: , np.newaxis], axis=1))

print("Top-1 Accuracy :", round(top1_accuracy, 1))
print("Top-2 Accuracy :", round(top2_accuracy, 1))
```

Gambar 4.17 Perhitungan *accuracy* pada kalimat kombinasi tiga aspek.

4.4 Proses Evaluasi Model Sentimen Aspek

Pada Gambar 4.18 merupakan langkah pembuatan *code* terakhir untuk menjalankan proses evaluasi pada sentimen kalimat review. Data *input* berupa file *validation_sentimen.xlsx* digunakan untuk validasi model sentimen. Data validasi sentimen terdiri dari beberapa kalimat yang mengandung kombinasi aspek. Pada tahap ini dilakukan proses cleansing data untuk memastikan data yang akan di proses sesuai dengan format model data.

```
import neattext.functions as nfx

data_validation = pd.read_excel('validation_sentimen.xlsx')
data_validation.tail()

def cleaning(text):
    text = re.sub(r'\n', ' ', text) # Hapus \n (enter)
    text = nfx.remove_hashtags(text) # Hapus hashtags
    text = nfx.remove_numbers(text) # Hapus number
    text = text.strip() # Hapus Whitespace
    text = re.sub('[^0-9a-zA-Z]+', ' ', text) # Hapus karakter selain alfabet dan angka
    return text

def casefolding(text):
    return text.lower()

data_validation['content_cleaned'] = data_validation['Text'].apply(cleaning).apply(casefolding)

data_validation.head()
```

Gambar 4.18 Inisiasi data validasi sentiment.

Model *sentiment_aspek.h5* merupakan model sentiment yang telah dibangun dengan LSTM. Data validasi diterapkan pada model untuk mengetahui hasil prediksi dari model yang dibangun dengan *output binary*. Hasil dari prediksi data disimpan kedalam *variable prediction_result*. Proses ini dapat dilihat pada Gambar 4.19.

```
from keras.models import load_model

model_filenames = ['sentiment_aspek.h5'] # List of models
prediction_sentiments = []
threshold = 0.5

models = []
for filename in model_filenames:
    model = load_model(filename)
    models.append(model)

predictions = np.zeros((test_pad.shape[0], len(model_filenames)))

for i, model in enumerate(models):
    preds = model.predict(test_pad)
    preds_binary = np.where(preds > 0.5, 1, 0)
    for j, pred in enumerate(preds_binary):
        predictions[j][i] = pred[0]

prediction_result = predictions.flatten().astype(int)
```

Gambar 4.19 Tahapan evaluasi sentimen.

Selanjutnya setelah melakukan proses prediksi dari data input, dilakukan proses perhitungan akurasi prediksi. Gambar 4.20 merupakan langkah penilaian akurasi prediksi. Penilaian akurasi model terhadap sentimen kalimat menggunakan F1_Score.

```
Accuracy = metrics.accuracy_score(actual_sentiment, prediction_result)
F1_score = metrics.f1_score(actual_sentiment, prediction_result)
print("Accuracy :", Accuracy)
print("F1_score :", F1_score)

Accuracy : 0.7647058823529411
F1_score : 0.8048780487804877
```

Gambar 4.20 Evaluasi hasil sentimen dengan F1_Score.

4.5 Pembahasan

4.5.1 Hasil Akurasi Training Model

Tahapan awal pada sentimen multi aspek yaitu pembangunan tujuh model. Enam model fokus pada aspek yang berbeda dalam ulasan hotel dan satu model untuk sentimen. Tabel 4.2 menunjukkan hasil akurasi training model pada tiap aspek dan sentimen kalimat.

Tabel 4.2 Hasil *Training* Model Aspek dengan Berbagai LSTM Unit

Unit LSTM	Harga	Lokasi	Pelayanan	Kamar	Fasilitas	Makanan	Sentimen
32	0.9247	0.9524	0.9150	0.9208	0.7031	0.9454	0.8477
64	0.9512	0.9643	0.9271	0.9472	0.8125	0.9538	0.8736
100	0.8703	0.7698	0.9231	0.7585	0.7031	0.9286	0.8160

Tabel 4.3 memberikan informasi yang sangat penting mengenai performa model selama proses pelatihan. Pengaturan tiga unit LSTM yang berbeda menghasilkan grafik yang beragam untuk masing-masing aspek model. Namun, secara keseluruhan, dapat disimpulkan bahwa model telah berfungsi dengan baik. Grafik menunjukkan bahwa garis *training loss* dan *validation loss* cenderung menurun seiring dengan berjalannya proses pelatihan. Hal ini menandakan bahwa model mampu belajar dari data training dengan baik dan mengurangi tingkat kesalahan pada data validasi. Selain itu, garis kurva *training accuracy* dan *validation accuracy* menunjukkan tren yang naik seiring dengan berjalannya pelatihan. Hal ini menunjukkan bahwa model mampu meningkatkan akurasi dalam memprediksi data training dan data validasi seiring dengan berjalannya waktu. Kinerja model yang baik ini memberikan keyakinan bahwa model siap digunakan untuk analisis lebih lanjut dan aplikasi praktis dalam mengenali aspek-aspek penting dalam ulasan pelanggan hotel.

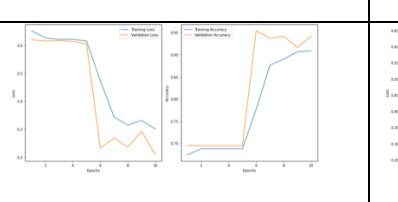
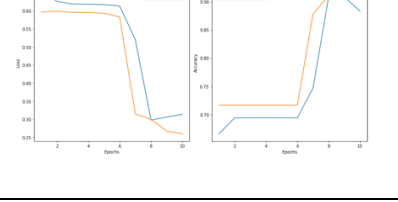
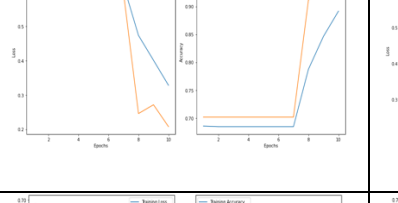
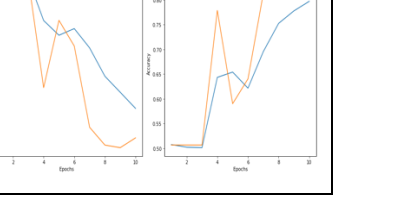
Tahapan training dengan pengaturan unit 32 dan 64, terjadi masalah *overfitting* pada aspek fasilitas dan harga. *Overfitting* terjadi ketika model terlalu rumit atau terlalu kompleks untuk data training yang terbatas, sehingga model "menghafal" data training dan tidak mampu menggeneralisasi dengan baik pada data yang belum pernah dilihat sebelumnya. Kondisi ini dapat mengakibatkan performa yang buruk dalam mengenali pola dan menghasilkan prediksi yang tidak akurat pada data uji.

Pengaturan unit 64, keseluruhan model aspek menunjukkan tren kinerja yang cukup baik. Grafik loss cenderung menurun seiring dengan berjalannya proses pelatihan, menandakan bahwa model mampu mengurangi tingkat kesalahan pada data training dan validasi. Selain itu, grafik accuracy menunjukkan tren yang naik secara bersamaan, menandakan bahwa model mampu meningkatkan akurasi dalam memprediksi data training dan data validasi.

Hasil ini menunjukkan bahwa model LSTM dengan pengaturan unit 64 mampu memberikan performa yang baik dalam memodelkan data teks dengan kompleksitas yang

berbeda. Model tidak terlalu sederhana sehingga tidak dapat menangkap pola yang rumit, dan juga tidak terlalu kompleks sehingga menyebabkan *overfitting*.

Tabel 4.3 Learning Curve Pada Setiap Training Model Aspek dengan LSTM Unit yang Berbeda

Aspek	32 unit	64 unit	100 unit
Harga			
Lokasi			
Pelayanan			
Kamar			
Fasilitas			
Makanan			
Sentimen			

Tahap pelatihan model ini, dilakukan berbagai percobaan dengan mengatur parameter yang berbeda, salah satunya adalah jumlah unit LSTM yang digunakan. Dalam eksperimen ini, tiga pengaturan jumlah unit LSTM diuji, yaitu 32, 64, dan 100. Dari hasil percobaan tersebut, ditemukan bahwa model dengan 64 unit LSTM memberikan akurasi terbaik dibandingkan dengan pengaturan lainnya. Meskipun pengaturan dengan 100 unit LSTM memiliki jumlah unit yang lebih besar, namun ternyata cenderung menghasilkan akurasi yang lebih rendah dibandingkan dengan 64 unit.

Sementara itu, pada pengaturan dengan 32 unit LSTM, masih terdapat hasil akurasi di bawah 80%, khususnya pada aspek fasilitas. Hasil ini menunjukkan bahwa penggunaan 64 unit LSTM memberikan keseimbangan yang baik antara kompleksitas model dan akurasi, dan dapat dijadikan pilihan yang lebih baik dalam membangun model yang efektif untuk analisis teks yang dilakukan. Selanjutnya, temuan ini dapat membantu dalam pengoptimalan parameter pada model dan meningkatkan performa model pada analisis multi aspek.

Pencapaian akurasi menunjukkan bahwa model-model yang dibangun berhasil dalam mengklasifikasikan ulasan hotel berdasarkan nilai yang dituju. Dengan akurasi yang baik, model tersebut menjadi acuan dalam identifikasi multi aspek. Sehingga model dapat memahami dan menganalisis setiap penilaian pada kalimat *review*.

4.5.2 Hasil Evaluasi Klasifikasi Aspek

Proses evaluasi klasifikasi multi aspek yang dilakukan dengan tiga model yang berbeda memiliki tujuan untuk memahami bagaimana performa klasifikasi berbeda-beda tergantung pada jumlah aspek yang diikutsertakan. Dalam hal ini, penulis memutuskan untuk membagi evaluasi menjadi tiga model yang berbeda:

1. Klasifikasi kalimat satu aspek: Model ini digunakan untuk melakukan klasifikasi sentimen pada satu aspek yang ada dalam review hotel. Misalnya, model ini dapat digunakan untuk mengklasifikasikan sentimen terkait dengan aspek pelayanan.
2. Klasifikasi kalimat dua aspek: Model ini digunakan untuk menggabungkan dua aspek yang ada dalam *review* hotel dan melakukan klasifikasi sentimen berdasarkan kombinasi tersebut. Contohnya, model ini dapat mengklasifikasikan sentimen terkait dengan aspek pelayanan dan fasilitas.
3. Klasifikasi kalimat tiga aspek: Model ini digunakan untuk menggabungkan tiga aspek yang ada dalam review hotel dan melakukan klasifikasi sentimen berdasarkan kombinasi tersebut. Contohnya, model ini dapat mengklasifikasikan sentimen terkait dengan aspek pelayanan, fasilitas, dan harga.

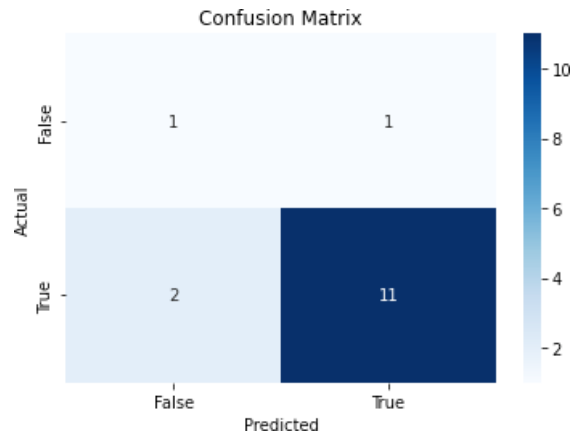
Pemilihan kombinasi maksimal tiga aspek didasarkan pada observasi bahwa rata-rata penulis review mengulas tidak lebih dari tiga aspek dalam satu review. Dengan menggunakan tiga model yang berbeda, penulis dapat memperoleh pemahaman yang lebih baik tentang kinerja klasifikasi dalam mengatasi jumlah aspek yang berbeda dalam review hotel.

Evaluasi pertama dilakukan dengan tujuan untuk menilai akurasi model dalam menghadapi kalimat yang memiliki satu aspek. Hasil evaluasi tersebut direkam dan ditampilkan dalam Tabel 4.4, yang memuat nilai F1 score, akurasi (*accuracy*), presisi (*precision*), dan recall untuk masing-masing model. F1 score merupakan ukuran gabungan dari presisi dan recall, yang memberikan informasi mengenai keseimbangan antara keduanya. Akurasi (*accuracy*) adalah persentase prediksi yang benar dari total data uji. Presisi (*precision*) mengukur sejauh mana hasil prediksi positif yang benar, sedangkan *recall* mengukur sejauh mana kemampuan model dalam menemukan kembali data positif yang sebenarnya. Hasil dari evaluasi ini memberikan wawasan terhadap performa model dalam mengklasifikasikan kalimat-kalimat dengan satu aspek.

Tabel 4.4 Hasil Evaluasi Kalimat Satu Aspek

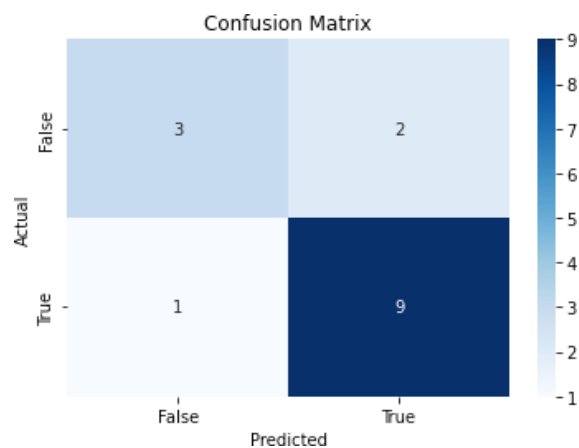
Aspek	F1_Score	Accuracy	Precision	Recall
Harga	75%	0.6	0.75	0.75
Pelayanan	76,1%	0.66	0.88	0.66
Makanan	87,9%	0.8	0.84	0.91
Fasilitas	78,2%	0.66	0.81	0.75
Lokasi	85,7%	0.8	0.9	0.818
Kamar	84,6%	0.73	0.78	0.91

Hasil evaluasi akurasi menunjukkan bahwa perolehan F1 score tertinggi terjadi pada aspek makanan, dengan nilai sebesar 87,9%. Untuk memvisualisasikan evaluasi aspek makanan, dapat dilihat pada Gambar 4.21 yang menampilkan confusion matrix. Dari seluruh data yang diberikan, terdapat tiga dari lima belas data yang diprediksi dengan ketidaktepatan oleh model. Hasil ini konsisten dengan akurasi model yang secara keseluruhan berada pada angka rata-rata 0,9 dari berbagai pengujian yang dilakukan. Hal ini menunjukkan bahwa model secara keseluruhan memiliki performa yang baik dan konsisten dalam mengklasifikasikan aspek makanan dengan tepat, meskipun masih terdapat beberapa data yang diprediksi dengan sedikit kesalahan.

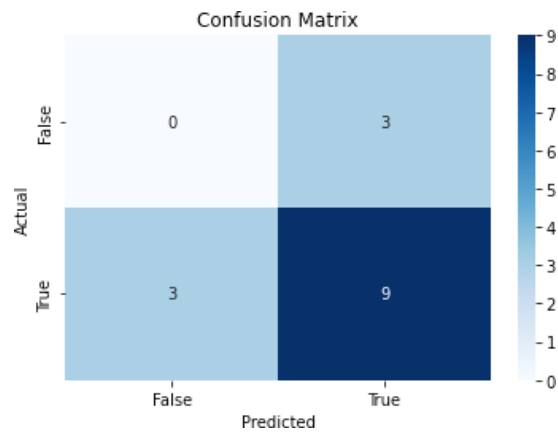


Gambar 4.21 Confusion matrix aspek makanan.

Evaluasi selanjutnya, hasil perolehan skor sebesar 84% dan 85% ditemukan pada aspek lokasi dan kamar. Dua gambar, yaitu Gambar 4.22 dan Gambar 4.23 menunjukkan hasil akurasi dari kedua aspek tersebut. Meskipun masih terdapat ketidaksesuaian dalam prediksi aspek tertentu, namun secara keseluruhan model telah mampu menebak aspek utamanya dengan akurasi yang baik. Hasil ini menunjukkan bahwa model dapat dengan cukup baik mengidentifikasi dan mengklasifikasikan ulasan terkait aspek lokasi dan kamar pada review hotel. Meskipun ada beberapa kasus di mana prediksi tidak sepenuhnya tepat, model secara keseluruhan telah memberikan hasil yang dapat diandalkan dalam mengenali aspek utama dalam setiap ulasan.

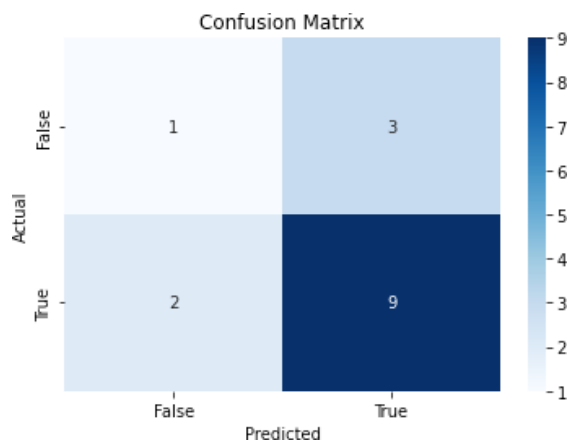


Gambar 4.22 Confusion matrix aspek lokasi.

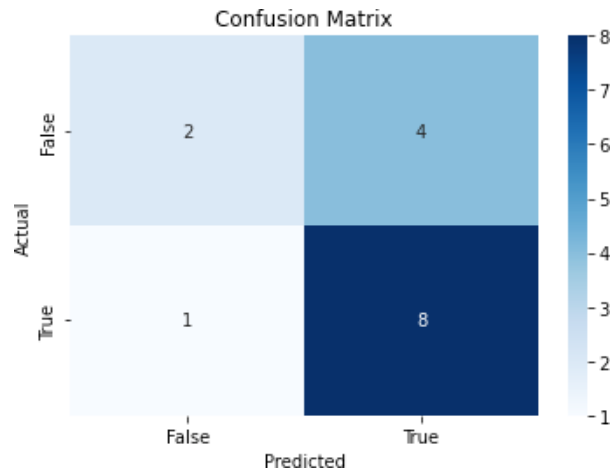


Gambar 4.23 Confusion matrix aspek kamar

Hasil evaluasi menunjukkan bahwa terdapat tiga aspek dengan akurasi terendah, yaitu fasilitas, pelayanan, dan harga. Ketiga aspek tersebut menunjukkan tingkat akurasi yang cukup rendah, dengan nilai di bawah 80%. Hasil ini mengindikasikan bahwa model menghadapi kesulitan dalam mengklasifikasikan ulasan terkait ketiga aspek tersebut, dan terdapat tantangan dalam mengenali dan memprediksi dengan tepat informasi terkait fasilitas, pelayanan, dan harga pada review hotel. Meskipun demikian, hasil ini tetap memberikan wawasan berharga tentang area yang memerlukan perbaikan dalam model dan analisis teks yang lebih mendalam pada aspek-aspek yang kritis bagi ulasan hotel.



Gambar 4.24 Confusion matrix aspek pelayanan



Gambar 4.25 Confusion matrix aspek harga

Gambar 4.26 merupakan salah satu hasil dari pengujian kalimat dengan satu aspek. Hasil prediksi berupa bilangan biner yang menyatakan bahwa angka 1 merupakan aspek utama, sementara angka 0 merupakan aspek lainnya. Pada Gambar 4.26 aspek yang diprediksi yaitu aspek makanan. *Actual_aspect* menunjukkan aspek utama yang sesungguhnya, sementara pada *predicted_aspect* merupakan hasil prediksi aspek dari pengujian model.

original_text	text_cleaned	actual_aspect	predicted_aspect
makanan yang disajikan kurang enak	makanan yang disajikan kurang enak	1	1
makanannya kurang enak, tapi overall good	makanannya kurang enak tapi overall good	1	1
Tempat makan nya agak sempit tapi makanan nya ...	tempat makan nya agak sempit tapi makanan nya ...	1	1
Makanannya enak dan bervariasi	makanannya enak dan bervariasi	1	1
sarapannya variatif dan enak	sarapannya variatif dan enak	1	1
Makanan enak & bervariasi. Rekomended hotel !!!	makanan enak bervariasi recomended hotel	1	1
menu makanan lengkap',	menu makanan lengkap	1	1
makanan dan minuman sarapan pagi enak enak',	makanan dan minuman sarapan pagi enak enak	1	1
makanan di hotel ini lezat dan memuaskan',	makanan di hotel ini lezat dan memuaskan	1	0
restoran hotel ini menyajikan makanan yang lez...	restoran hotel ini menyajikan makanan yang lez...	1	1
Makanan di hotel ini sungguh memuaskan selera',	makanan di hotel ini sungguh memuaskan selera	1	0
aneh banget	aneh banget	0	0
ini rumah horor	ini rumah horor	0	1
tidak suka permainan ini	tidak suka permainan ini	0	0
pelayanan seperti bintang lima,	pelayanan seperti bintang lima	0	0
pelayanan sopan ramah	pelayanan sopan ramah	0	1
pelayanan cekatan',	pelayanan cekatan	0	1

Gambar 4.26 Hasil prediksi kalimat dengan aspek makanan.

Dalam penelitian ini, dilakukan pengujian terhadap total 30 kombinasi aspek yang terbentuk dari kombinasi dua aspek. Hasil evaluasi dari pengujian kombinasi dua aspek tersebut terdokumentasi dalam Tabel 4.5. Dari hasil pengujian tersebut, ditemukan bahwa skor tertinggi sebesar 79% berhasil dicapai pada kombinasi aspek fasilitas dan makanan. Hasil ini menunjukkan bahwa model secara efektif dapat mengklasifikasikan dan memprediksi ulasan yang berkaitan dengan aspek fasilitas dan makanan dengan akurasi yang tinggi. Meskipun ada beberapa kombinasi aspek lain yang juga menunjukkan performa yang baik, namun perolehan skor tertinggi pada kombinasi fasilitas - makanan mengindikasikan bahwa kedua aspek ini memiliki hubungan yang penting dalam ulasan hotel dan dapat dengan baik diidentifikasi oleh model.

Tabel 4.5 Hasil Evaluasi Kalimat Kombinasi Dua Aspek

Kombinasi 2 Aspek	Akurasi kombinasi 2-aspek
Fasilitas - Kamar	79%
Fasilitas - Makanan	70%
Makanan - Harga	67%
Makanan - Pelayanan	66%
Harga - Lokasi	65%
Harga - Pelayanan	66%
Makanan – Kamar	69%
Harga - Makanan	67%
Pelayanan - Makanan	69%
Kamar - Fasilitas	66%
Lokasi - Fasilitas	65%
Pelayanan - Fasilitas	69%
Harga - Kamar	72%
Lokasi - Kamar	69%
Pelayanan - Kamar	70%
Lokasi - Harga	70%
Pelayanan - Harga	70%
Pelayanan - Lokasi	70%
Makanan - Fasilitas	60%
Makanan - Lokasi	60%
Fasilitas - Harga	60%
Fasilitas - Lokasi	60%
Kamar - Harga	60%
Kamar - Lokasi	60%
Lokasi - Pelayanan	60%
Kamar - Makanan	61%

Lokasi - Makanan	62%
Harga - Fasilitas	61%
Kamar - Pelayanan	45%
Fasilitas - Pelayanan	41%

Evaluasi model dengan kombinasi kalimat dua aspek dapat terlihat hasilnya pada Gambar 4.27. Model mampu melakukan prediksi dengan tepat pada kalimat `Saya sangat puas dengan pelayanan yang diberikan di restoran ini. Pelayanannya cepat dan efisien, dan para stafnya sangat ramah. Makanannya juga luar biasa, rasanya enak dan terasa segar.`. Namun secara keseluruhan model hanya mampu melakukan prediksi akurat pada salah satu aspek saja. Hal ini mengindikasikan bahwa ketepatan model dalam memprediksi secara akurat pada kedua aspek kombinasi dalam satu rangkaian kalimat masih terbilang kurang baik.

original_text	text_cleaned	actual_1	actual_2	predicted_1	predicted_2
Kamar di hotel ini sangat nyaman dengan fasilitas l...	Kamar di hotel ini sangat nyaman dengan fasilitas...	2.0	4.0	0	0
Saya sangat puas dengan kamar yang disediakan ...	Saya sangat puas dengan kamar yang disediakan...	2.0	4.0	1	1
Kamar yang luas dan fasilitas yang lengkap memb...	Kamar yang luas dan fasilitas yang lengkap mem...	2.0	4.0	0	1
Harga yang terjangkau dengan fasilitas lengkap me...	Harga yang terjangkau dengan fasilitas lengkap m...	2.0	0.0	0	0
Saya sangat senang dengan harga yang kompetitif...	Saya sangat senang dengan harga yang kompetit...	2.0	0.0	1	1
Fasilitas yang sangat baik dengan harga yang terja...	Fasilitas yang sangat baik dengan harga yang terj...	2.0	0.0	1	1
Lokasi hotel ini sangat strategis dengan akses mud...	Lokasi hotel ini sangat strategis dengan akses mu...	2.0	0.0	1	1
Saya sangat puas dengan lokasi hotel ini yang ber...	Saya sangat puas dengan lokasi hotel ini yang be...	2.0	0.0	1	1
Lokasi hotel ini sangat nyaman dan tenang, jauh d...	Lokasi hotel ini sangat nyaman dan tenang jauh d...	2.0	0.0	1	1
Pelayanan yang ramah dan fasilitas yang lengkap ...	Pelayanan yang ramah dan fasilitas yang lengkap...	2.0	0.0	1	1
Kami sangat terkesan dengan pelayanan yang ram...	Kami sangat terkesan dengan pelayanan yang ra...	2.0	0.0	0	0
Pelayanan yang sangat baik dan fasilitas yang leng...	Pelayanan yang sangat baik dan fasilitas yang len...	2.0	0.0	0	1
Harga kamar di hotel ini sangat terjangkau, membu...	Harga kamar di hotel ini sangat terjangkau memb...	0.0	4.0	0	0
Saya sangat terkesan dengan kualitas kamar yang ...	Saya sangat terkesan dengan kualitas kamar yan...	0.0	4.0	1	1
Harga kamar yang ditawarkan sangat sebanding d...	Harga kamar yang ditawarkan sangat sebanding ...	0.0	4.0	0	0
Lokasi hotel ini sangat strategis dan kamar yang di...	Lokasi hotel ini sangat strategis dan kamar yang d...	0.0	4.0	1	1

Gambar 4.27 Hasil prediksi model dengan kombinasi kalimat dua aspek.

Terdapat total 60 kombinasi yang berbeda untuk evaluasi kombinasi 3 aspek. Akurasi kombinasi 3-aspek dan 2-aspek digunakan untuk mengukur kinerja model dalam mengklasifikasikan sentimen pada kombinasi aspek yang ada. Hasil evaluasi dapat terlihat pada

Tabel 4.6.

Tabel 4.6 Hasil Evaluasi Kalimat Kombinasi Tiga Aspek

Kombinasi 3 Aspek	Akurasi kombinasi 3-aspek	Akurasi kombinasi 2-aspek
Harga - Kamar - Pelayanan	53%	79%
Lokasi - Kamar - Makanan	42%	79%
Harga - Kamar - Makanan	52%	78%
Harga - Fasilitas - Pelayanan	53%	74%
Harga - Lokasi - Makanan	58%	68%
Harga - Makanan - Lokasi	53%	68%
Harga - Makanan - Lokasi	53%	68%
Fasilitas - Makanan - Kamar	68%	63%
Fasilitas - Makanan - Kamar	68%	63%
Harga - Fasilitas - Makanan	58%	63%
Harga - Makanan - Fasilitas	53%	63%
Harga - Lokasi - Pelayanan	53%	63%
Harga - Fasilitas - Lokasi	52%	63%
Harga - Makanan - Pelayanan	47%	63%
Kamar - Lokasi - Pelayanan	47%	63%
Lokasi - Makanan - Pelayanan	32%	63%
Lokasi - Harga - Makanan	31%	63%
Fasilitas - Kamar - Pelayanan	58%	58%
Kamar - Harga - Makanan	47%	58%
Makanan - Lokasi - Pelayanan	42%	58%
Lokasi - Makanan - Fasilitas	32%	58%
Lokasi - Pelayanan - Makanan	21%	58%
Makanan - Kamar - Lokasi	68%	53%
Makanan - Harga - Pelayanan	58%	53%
Kamar - Fasilitas - Lokasi	47%	53%
Kamar - Makanan - Lokasi	42%	53%
Kamar - Fasilitas - Pelayanan	37%	53%
Kamar - Makanan - Pelayanan	37%	53%
Harga - Kamar - Lokasi	52%	52%
Kamar - Pelayanan - Fasilitas	42%	52%
Fasilitas - Harga - Pelayanan	63%	47%
Kamar - Harga - Lokasi	58%	42%

Hasil evaluasi menunjukkan bahwa terdapat dua kombinasi aspek yang memiliki kombinasi 3-aspek accuracy tertinggi, yaitu harga-fasilitas-lokasi dan kamar-pelayanan-

fasilitas, dengan tingkat keakuratan sebesar 63%. Selanjutnya kombinasi kedua aspek tersebut memperoleh kombinasi 2-aspek accuracy tertinggi sebesar 79%. kombinasi 3-aspek akurasi menggambarkan kemampuan model untuk secara tepat memprediksi tiga kombinasi aspek secara keseluruhan. Sementara itu kombinasi 2-aspek akurasi memberikan fleksibilitas dengan menganggap prediksi sudah benar asalkan dua dari tiga aspek yang seharusnya diprediksi benar telah diidentifikasi dengan tepat. Hasil evaluasi ini menunjukkan tingkat kinerja model dalam menghadapi berbagai kombinasi aspek secara komprehensif.

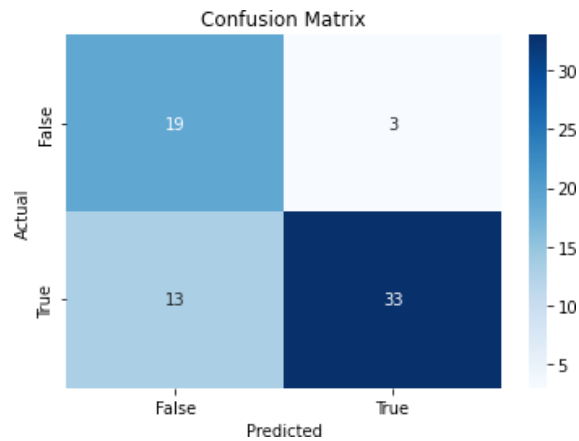
Dari daftar hasil prediksi, terlihat bahwa beberapa nilai akurasi kombinasi 2-aspek masih lebih rendah daripada kombinasi 3-aspek. Hal ini terjadi karena beberapa model hanya mampu memprediksi satu dari tiga aspek dengan benar. Ketidakmampuan model untuk mengidentifikasi dua aspek dari tiga aspek yang seharusnya diprediksi secara tepat menyebabkan hasil akurasi 2-aspek menjadi relatif lebih rendah.

Kondisi ini menyoroti pentingnya untuk meningkatkan kemampuan model dalam mengenali lebih banyak aspek dengan benar. Dengan demikian, performa model dapat ditingkatkan dan hasil akurasi 2-aspek dapat mendekati atau bahkan melampaui akurasi 3-aspek, menunjukkan kemampuan yang lebih baik dalam memprediksi kombinasi aspek secara tepat.

Kombinasi aspek harga-fasilitas-lokasi adalah kombinasi yang mencakup faktor harga, fasilitas, dan lokasi hotel. Dalam penilaian pelanggan, kombinasi ini mungkin menjadi pertimbangan utama dalam mengukur kualitas dan nilai dari sebuah hotel. Sementara itu, kombinasi aspek kamar-pelayanan-fasilitas melibatkan penilaian terhadap kualitas kamar, pelayanan, dan fasilitas yang tersedia di hotel. Sentimen yang diungkapkan dalam kombinasi ini juga memiliki tingkat akurasi yang tinggi, menunjukkan pentingnya faktor-faktor ini dalam mempengaruhi pengalaman dan kepuasan pelanggan.

4.5.3 Hasil Evaluasi Sentimen

Hasil perhitungan matriks konfusi terhadap prediksi sentiment terlihat pada Gambar 4.28. Dari total 68 kalimat review didapatkan nilai TF = 13, TT = 33, FT = 3, dan FF = 19. Dalam pengujian sentiment kalimat masih ditemukan ketidaksesuaian saat melakukan prediksi label. Kalimat review dengan kombinasi aspek lebih dari satu cenderung menghasilkan prediksi yang kurang tepat.



Gambar 4.28 Matriks hasil evaluasi sentiment kalimat.

Gambar 4.29 menunjukkan hasil prediksi sentimen dari kalimat review. Meskipun hasil prediksi secara keseluruhan menunjukkan kinerja yang baik, namun terdapat beberapa kalimat yang mengandung kombinasi aspek yang masih mengalami kesalahan prediksi. Sebagai contoh, pada kalimat 'Tempat Oke. Tapi untuk makanan sangat sangat buruk,' model masih memprediksi kalimat tersebut sebagai kalimat positif, padahal dari aspek lokasi, kalimat tersebut seharusnya memiliki sentimen yang lebih negatif.

original_text	text_cleaned	actual_sentiment	predicted_sentiment
kamar yg bersih fasilitas oje lokasi dekat dg penjual makanan serta pelayanan yg r...	kamar yg bersih fasilitas oje lokasi dekat dg penjual makanan serta pelayanan ...	1	0
respon staff baik dan cepat',	respon staff baik dan cepat	1	1
pelayanan semua ramah dan bagus saya suka',	pelayanan semua ramah dan bagus saya suka	1	1
fasilitas hotel bagus yang sangat lengkap seperti bintang lima',	fasilitas hotel bagus yang sangat lengkap seperti bintang lima	1	1
saya suka lokasi hotel strategis dekat banyak rumah makan',	saya suka lokasi hotel strategis dekat banyak rumah makan	1	1
kamarnya juga bersih dan nyaman',	kamarnya juga bersih dan nyaman	1	1
makanan dan minuman untuk sarapan pagi bervariasi buat anak-anak suka',	makanan dan minuman untuk sarapan pagi bervariasi buat anak anak suka	1	1
makanan dan minuman sarapan pagi enak semua',	makanan dan minuman sarapan pagi enak semua	1	1
makanan di hotel ini lezat',	makanan di hotel ini lezat	1	1
kamar mewah dengan balkon pemandangan kota',	kamar mewah dengan balkon pemandangan kota	1	1
harga terjangkau kualitas bintang lima, saya sangat suka',	harga terjangkau kualitas bintang lima saya sangat suka	1	1
Kurang bersih menurut saya, pelayanan kurang ok, karena saya ditelpon tengah m...	kurang bersih menurut saya pelayanan kurang ok karena saya ditelpon tengah ...	0	0
Pelayanan tdk menyenangkan. Kamar ada keacoaknya.	pelayanan tdk menyenangkan kamar ada keacoaknya	0	1
Kamar tidak kedap suara. Sangat mengganggu kenyamanan	kamar tidak kedap suara sangat mengganggu kenyamanan	0	0
Kamar nyaman tapi tidak kedap suara	kamar nyaman tapi tidak kedap suara	0	0
Tempat Oke Tapi untuk makanan sangat sangat buruk	tempat oke tapi untuk makanan sangat sangat buruk	0	1
Ukuran kamar kecil, shower tidak berfungsi dgn baik	ukuran kamar kecil shower tidak berfungsi dgn baik	0	0
Ruangan pengap. Pelayanan kurang.	ruangan pengap pelayanan kurang	0	1
kebersihan kamarnya tidak dijaga, shower rusak tidak diganti, dinding sampai jamu...	kebersihan kamarnya tidak dijaga shower rusak tidak diganti dinding sampai ja...	0	0
harga rada mahal untuk sekelas POP Hotel	harga rada mahal untuk sekelas pop hotel	0	1

Gambar 4.29 Hasil prediksi sentimen review.

Setelah melakukan proses prediksi kalimat dan melakukan perhitungan akurasi dengan F1_Score, hasil evaluasi menunjukkan bahwa model memiliki akurasi sebesar 76% dan F1_Score sebesar 83%. Nilai tersebut terbilang cukup baik, menandakan bahwa model mampu memberikan prediksi yang relatif akurat dan konsisten terhadap sentimen kalimat review.

BAB 5

Kesimpulan dan Saran

5.1 Kesimpulan

Dalam penelitian ini, telah berhasil dilakukan pembuatan model sentimen multi aspek pada data review hotel dengan menggunakan metode LSTM. Terdapat 7 model yang terdiri dari 6 model aspek dan 1 model sentimen. Seluruh model tersebut dapat menggambarkan nilai-nilai yang terkandung dalam kalimat review, terlepas dari kombinasi aspek yang ada.

Hasil pengujian pada model multi aspek menunjukkan bahwa performa terbaik diperoleh pada pengujian kalimat dengan satu aspek dengan F1 score sebesar 85,7%. Namun, untuk kalimat kombinasi dua dan tiga aspek, akurasi yang dicapai relative lebih rendah yaitu sebesar 79%. Hal ini menunjukkan bahwa model dapat dengan mudah mengidentifikasi aspek dalam kalimat satu aspek, tetapi terkendala dalam mendeteksi keseluruhan aspek dalam kalimat multi aspek.

Selanjutnya pada evaluasi model sentimen, penelitian ini telah mencapai performa terbaik dengan akurasi sebesar 76% dan nilai F1 Score mencapai 83%. F1 Score adalah metrik gabungan dari precision dan recall yang memberikan gambaran keseluruhan tentang kemampuan model dalam memprediksi sentimen positif dan negatif kalimat review. Meskipun begitu, perlu diperhatikan bahwa model sentimen ini fokus pada sentimen keseluruhan dan belum mempertimbangkan aspek-aspek khusus dalam kalimat yang bisa menjadi perbaikan berikutnya.

Pemodelan multi aspek ini memungkinkan kita untuk mengetahui dengan lebih spesifik aspek-aspek apa saja yang terdapat dalam kalimat review. Hal ini memberikan manfaat yang lebih baik dalam memahami penilaian hotel yang terkandung dalam kalimat review tersebut. Dengan demikian, model ini dapat dimanfaatkan secara lebih efektif untuk analisis sentimen multi aspek pada review hotel.

5.2 Saran

Dari hasil penelitian terdapat dua saran yang dapat digunakan sebagai pengembangan selanjutnya, yaitu:

1. Perlu dilakukan pengembangan lebih lanjut untuk meningkatkan kemampuan model dalam mengenali dan menggambarkan aspek-aspek pada kalimat dengan kombinasi aspek yang lebih banyak.

2. Perlu melakukan pengembangan untuk mengetahui keterkaitan antara sentimen dan setiap aspek penilaian dalam kalimat review.
3. Perlu dilakukan uji coba dengan teknik metode deep learning yang lain untuk memperluas variasi dan kuantitas data yang tersedia.

Daftar Pustaka

- Aminullah, M. (n.d.). *Alat Deteksi Masker Dengan Metode Convolutional Neural Network Untuk Tunanetra Pada Era New Normal*.
- Bouazizi, M., & Ohtsuki, T. (2017). A Pattern-Based Approach for Multi-Class Sentiment Analysis in Twitter. *IEEE Access*, 5, 20617–20639.
<https://doi.org/10.1109/ACCESS.2017.2740982>
- Cahyaningtyas, S., Hatta Fudholi, D., & Fathan Hidayatullah, A. (2021). Deep Learning for Aspect-Based Sentiment Analysis on Indonesian Hotels Reviews. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*. <https://doi.org/10.22219/kinetik.v6i3.1300>
- Haque, R., Islam, N., Tasneem, M., & Das, A. K. (2023). Multi-class sentiment classification on Bengali social media comments using machine learning. *International Journal of Cognitive Computing in Engineering*, 4, 21–35.
<https://doi.org/10.1016/j.ijcce.2023.01.001>
- Hoang, M., & Rouces, J. (2019). *Aspect-Based Sentiment Analysis Using BERT*.
- Hochreiter, sepp, & Schmidhuber, J. (n.d.). *LONG SHORT-TERM MEMORY*.
- Hu, M., Zhao, S., Zhang, L., Cai, K., Su, Z., Cheng, R., & Shen, X. (2018). *CAN: Constrained Attention Networks for Multi-Aspect Sentiment Analysis*.
<http://arxiv.org/abs/1812.10735>
- Jayanto, R., Kusumaningrum, R., & Wibowo, A. (2022). Aspect-based sentiment analysis for hotel reviews using an improved model of long short-term memory. *International Journal of Advances in Intelligent Informatics*, 8(3), 391–403.
<https://doi.org/10.26555/ijain.v8i3.691>
- Liu, N., Shen, B., Zhang, Z., Zhang, Z., & Mi, K. (2019). Attention-based Sentiment Reasoner for aspect-based sentiment analysis. *Human-Centric Computing and Information Sciences*, 9(1), 35. <https://doi.org/10.1186/s13673-019-0196-3>
- Ma, X., Zeng, J., Peng, L., Fortino, G., & Zhang, Y. (2019). Modeling multi-aspects within one opinionated sentence simultaneously for aspect-level sentiment analysis. *Future Generation Computer Systems*, 93, 304–311.
<https://doi.org/10.1016/J.FUTURE.2018.10.041>

- Maksun, A. M., Sari, Y. A., & Rahayudi, B. (2021). *Analisis Sentimen pada Twitter Bencana Alam di Kalimantan Selatan menggunakan Metode Naïve Bayes* (Vol. 5, Issue 12). <http://j-ptiik.ub.ac.id>
- Malik, R. A. A., & Sibaroni, Y. (2022). Multi-aspect Sentiment Analysis of Tiktok Application Usage Using FasText Feature Expansion and CNN Method. *Journal of Computer System and Informatics (JoSYC)*, 3(4), 277–285. <https://doi.org/10.47065/josyc.v3i4.2033>
- Mazari, A. C., Boudoukhani, N., & Djeflal, A. (2023). BERT-based ensemble learning for multi-aspect hate speech detection. *Cluster Computing*, 1–15. <https://doi.org/10.1007/S10586-022-03956-X/METRICS>
- Mubarok, M. S., Adiwijaya, A., & Aldhi, M. D. (2017). Aspect-based sentiment analysis to review products using Naïve Bayes. *AIP Conference Proceedings*, 1867. <https://doi.org/10.1063/1.4994463>
- Priyantina, R. A., & Sarno, R. (2019). Sentiment analysis of hotel reviews using Latent Dirichlet Allocation, semantic similarity and LSTM. *International Journal of Intelligent Engineering and Systems*, 12(4), 142–155. <https://doi.org/10.22266/ijies2019.0831.14>
- Program, E. I., Sistem, S., Kampus, I. A., & Bogor, K. (2019). *Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes*. VII(1).
- Simonyan, K., & Zisserman, A. (2015). *VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION*. <http://www.robots.ox.ac.uk/>
- Tran, T., Ba, H., & Huynh, V. N. (2019). Measuring hotel review sentiment: An aspect-based sentiment analysis approach. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11471 LNAI, 393–405. https://doi.org/10.1007/978-3-030-14815-7_33
- Utami, E. (2023). OPTIMIZING SENTIMENT ANALYSIS OF PRODUCT REVIEWS ON MARKETPLACE USING A COMBINATION OF PREPROCESSING TECHNIQUES, WORD2VEC, AND CONVOLUTIONAL NEURAL NETWORK OPTIMISASI ANALISIS SENTIMEN ULASAN PRODUK PADA MARKETPLACE DENGAN KOMBINASI TEKNIK PREPROCESSING, WORD2VEC, DAN CONVOLUTIONAL NEURAL NETWORK. *Jurnal Teknik Informatika (JUTIF)*, 4(1), 101–107. <https://doi.org/10.20884/1.jutif.2023.4.1.815>

- Veronika Aritonang, Y., Purnama Napitupulu, D., Halomoan Sinaga, M., Amalia, J., Industri dan Teknik Elektro, F., & Teknologi Del, I. (2022). Pengaruh Hyperparameter Pada Fasttext Terhadap Performa Model Deteksi Sarkasme Berbasis Bi-LSTM. *Jurnal Teknik Informatika Dan Sistem Informasi*, 9(3).
<http://jurnal.mdp.ac.id>
- Widayat, W. (2021). Analisis Sentimen Movie Review menggunakan Word2Vec dan metode LSTM Deep Learning. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(3), 1018. <https://doi.org/10.30865/mib.v5i3.3111>
- Yutika, C. H., Adiwijaya, A., & Faraby, S. Al. (2021). Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2), 422.
<https://doi.org/10.30865/mib.v5i2.2845>
- Zaman, L., Sumpeno, S., & Hariadi, M. (2019). Analisis Kinerja LSTM dan GRU sebagai Model Generatif untuk Tari Remo. In *JNTETI* (Vol. 8, Issue 2).

LAMPIRAN

Tabel lampiran berisi proses training, dan testing beserta pengujian model.

<https://github.com/windiastri/multi-aspect-analysis>

Model dan Data	Link github
Training model aspek	https://github.com/windiastri/multi-aspect-analysis/blob/main/multi-aspect-v2.ipynb
Training model sentimen	https://github.com/windiastri/multi-aspect-analysis/blob/main/sentiment_aspek.ipynb
Validasi kalimat satu aspek	https://github.com/windiastri/multi-aspect-analysis/blob/main/validate-single-aspect.ipynb
Validasi kalimat kombinasi dua aspek	https://github.com/windiastri/multi-aspect-analysis/blob/main/validate-double-aspect.ipynb
Validasi kalimat kombinasi tiga aspek	https://github.com/windiastri/multi-aspect-analysis/blob/main/validate-triple-aspects.ipynb
Model Aspek	https://github.com/windiastri/multi-aspect-analysis/tree/main/data_training
Data training	https://github.com/windiastri/multi-aspect-analysis/tree/main/data_training
Data evaluasi	https://github.com/windiastri/multi-aspect-analysis/tree/main/data_evaluation