**Terakreditasi SINTA Peringkat 3**
Surat Keputusan Direktur Jenderal Pendidikan Tinggi, Riset, dan Teknologi Nomor 225/E/KPT/2022
masa berlaku mulai Vol.7 No. 1 tahun 2022 s.d Vol. 11 No. 2 tahun 2026

Terbit online pada laman web jurnal:
http://publishing-widyagama.ac.id/ejournal-v2/index.php/jointecs

JOINTECS

**JOINTECS**
**(Journal of Information Technology and Computer Science)**

Vol. 8 No. 2 (2023) 67 - 76

e-ISSN:2541-6448 p-ISSN:2541-3619

# Analisis Sentimen Berbasis Aspek Ulasan Pelanggan Restoran Menggunakan LSTM Dengan Adam *Optimizer*

Wardianto[1], Farikhin[2], Dinar Mutiara Kusumo Nugraheni[3]
[1]Magister Sistem Informasi, Sekolah Pascasarjana, Universitas Diponegoro Semarang
[2]Program Studi Matematika, Fakultas Sains dan Matematika, Universitas Diponegoro Semarang
[3]Program Studi Teknik Informatika, Fakultas Sains dan Matematika, Universitas Diponegoro Semarang
[1]wardianto@students.undip.ac.id, [2]farikhin@lecturer.undip.ac.id, [3]dinar.mutiara@live.undip.ac.id

***Abstract***
*Consumers believe that restaurant reviews are very important when choosing a restaurant. Due to the fact that reviews have become one of the most effective ways to influence customer decisions, research that has been done on restaurant customer reviews is about sentiment analysis. Previous studies have only used sentiment analysis at the sentence or document level, while a better level uses Aspect-Based Sentiment Analysis (ABSA), or a type of aspect-based sentiment analysis. LSTM is a variant of RNN that stores long-term information in memory cells. Use of global max pooling to reduce output resolution features and prevent overfitting. In addition, the optimization method used by Adam Optimizer is an adaptive learning rate optimization algorithm specifically designed to train deep neural networks. This study aims to classify restaurant customer opinions based on aspects (food, place, service, and price) based on restaurant customer reviews on Indonesian-language TripAdvisor with LSTM and global max pooling for sentiment classification (negative, half negative, neutral, half positive, positive). The results of this study indicate that the ABSA in restaurant customer reviews for sentiment classification accuracy is 78.7% and the aspect category accuracy is 78%, both are interconnected and can help understand restaurant customer opinions on TripAdvisor.*

*Keywords*: *Tripadvisor; LSTM; ABSA; Adam Optimizer*.

**Abstrak**

Konsumen mengatakan ulasan tentang restoran merupakan faktor penting saat memilih tempat makan, hal ini karena ulasan telah menjadi salah satu alat yang paling kuat untuk mempengaruhi keputusan para konsumen. *Deep learning* adalah sub bidang *machine learning* yang menggunakan jaringan saraf untuk memecahkan masalah yang sangat kompleks, LSTM adalah salah satu varian *RNN* yang menyimpan informasi jangka panjang ke dalam sel memori dan mempelajari informasi kontekstual yang berguna untuk tugas klasifikasi. Dan penggunaan *Global max pooling* untuk mengurangi fitur resolusi *output* dan mencegah *overfitting*. Selain itu metode optimasi yang digunakan adalah *Adam Optimizer* Yang merupakan algoritme pengoptimalan laju pembelajaran adaptif yang dirancang khusus untuk melatih *deep neural network.* Penelitian ini bertujuan untuk melakukan klasifikasi opini pelanggan restoran berbasis aspek (makanan, tempat, pelayanan, harga), berdasarkan ulasan pelanggan restoran pada *TripAdvisor* berbahasa Indonesia menggunakan metode LSTM dan *Global max pooling* untuk klasifikasi sentimen (negatif, setengah negatif, netral, setengah positif, positif). Hasil dari penelitian ini menunjukkan bahwa analisis sentimen berbasis aspek untuk ulasan pelanggan sangat penting dalam mengetahui klasifikasi sentimen dengan akurasi mencapai 78.7% dan kategori aspek memiliki akurasi 78%, kedua saling berhubungan dan dapat membantu memahami opini pelanggan restoran pada *TripAdvisor*.

Kata Kunci: *Tripadvisor; LSTM; ABSA; Adam Optimizer*.

CC BY SA

Diterima Redaksi : 12-06-2023 | Selesai Revisi : 26-06-2023 | Diterbitkan Online : 01-07-2023

## 1. Pendahuluan

Membaca ulasan konsumen sebelum membeli telah menjadi kebiasaan bagi banyak konsumen[1], hal ini karena ulasan telah menjadi salah satu alat yang paling kuat untuk mempengaruhi keputusan para konsumen [2]. Dari 91% konsumen mengatakan ulasan tentang restoran merupakan faktor penting saat memilih tempat makan[3]. Ulasan tidak hanya dapat digunakan oleh konsumen untuk mengekspresikan sikap dan sentimen mereka, tetapi juga dapat digunakan oleh pemilik usaha untuk mengamati reaksi konsumen dan mengusulkan sarana perbaikan[4].

Dengan mempertahankan tingkat kepuasan pelanggan sangat penting untuk menjaga bisnis restoran dengan baik[5], karena tujuan yang ideal dari restoran di antaranya menyenangkan pelanggan[6]. Bagi manajer restoran, ulasan positif yang diposting oleh pelanggan restoran dapat digunakan untuk meningkatkan reputasi atau kualitas pada restoran itu sendiri[7]. *TripAdvisor* merupakan situs web perjalanan terbesar didunia, yang menyajikan berbagai informasi kepada wisatawan termasuk di dalamnya informasi restoran[8].

Menurut statistik *TripAdvisor*, jumlah ulasan dan opini pengguna di *TripAdvisor* telah meningkat pesat dari 200 juta menjadi 1,000 juta dari 2014 hingga 2021[3]. Pengguna *TripAdvisor* memberikan ulasan pengalaman mereka dan memberikan peringkat menggunakan skor mulai dari 1 sampai 5, yang mewakili dari yang terburuk hingga yang sangat baik. Penilaian ini dianggap tidak relevan karena sulit mendapatkan nilai pasti yang diulas pengunjung di bagian ulasan[9]. Karena sebenarnya pengunjung menulis kalimat negatif meskipun memberi nilai 4 dan 5[10]. Oleh karena itu, pemilik restoran perlu mempertimbangkan umpan balik pengunjung untuk perbaikan dimasa depan guna menarik perhatian pelanggan[11].

Untuk memahami ulasan, dibutuhkan metode Analisis sentimen atau *opinion mining* yang merupakan cabang *Natural Language Processing (NLP)* untuk mengukur dan memprediksi sentimen pengguna[12], Mengingat perkembangan web 2.0 dan jejaring sosial berkembang pesat[10]. Secara tradisional analisis sentimen mengevaluasi ulasan berdasarkan kalimat atau dokumen[13], sedangkan tingkatan yang lebih baik menggunakan *Aspect-Based Sentiment Analysis (ABSA)* atau jenis analisis sentimen berbasis aspek[14].

*Deep learning* adalah sub bidang *machine learning* yang menggunakan jaringan saraf untuk memecahkan masalah yang sangat kompleks, Berkembangnya *deep learning* yang dapat digunakan untuk analisis sentimen yang memiliki kinerja lebih baik dari *machine learning* [15]. *Long Short-Term Memory* (LSTM) adalah salah satu varian *Recurrent Neural Network (RNN)* yang menyimpan informasi jangka panjang ke dalam sel memori dan mempelajari informasi kontekstual yang berguna untuk tugas klasifikasi[16]. Jaringan LSTM dirancang dengan tiga unit yaitu gerbang input, gerbang keluaran, dan gerbang lupa untuk mengontrol keadaan interval dan keluaran setiap saat, akibatnya LSTM dapat mengingat dengan lebih baik dalam urutan Panjang dan menangani dependensi jangka Panjang dengan menentukan informasi yang dibuang[17].

Pada penelitian sebelumnya, tentang ulasan restoran oleh [11], [18], [19], dan [20], Keterbatasan pada penelitian sebelumnya penggunaan analisis sentimen yang dilakukan pada tingkat kalimat atau tingkat dokumen, tidak melakukan kalsifikasi sentimen berdasarkan aspek yang ada dalam ulasan[21]. Dalam hal ini perlu dilakukan analisis sentimen tingkat lanjut seperti penggunaan *Aspect-Based Sentiment Analysis*(ABSA) tidak hanya mengumpulkan sentimen yang diekspresikan terhadap setiap aspek dalam ulasan pelanggan. Akan tetapi juga membantu mengekstrak aspek penting dari ulasan karena mengetahui hanya hal-hal positif dan negatif saja tidak cukup[22].

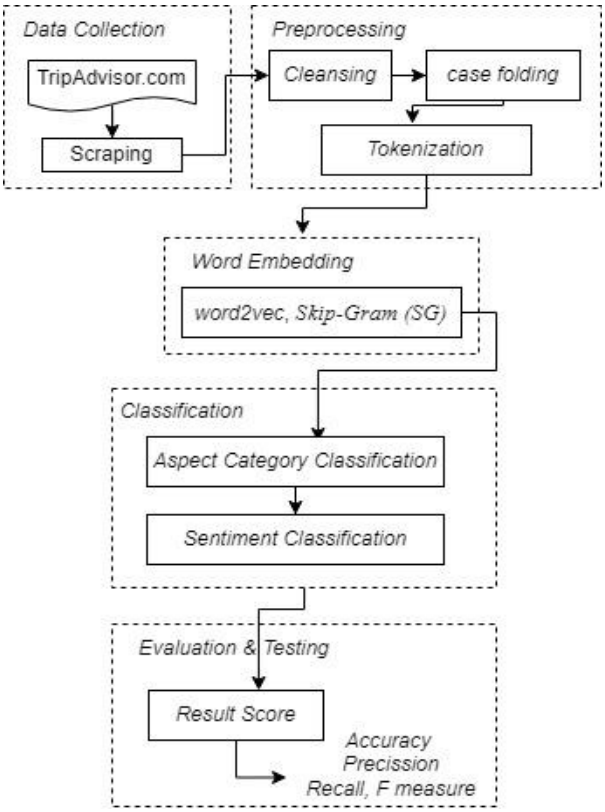
Berkembangnya Algoritma *deep learning* memberikan kesempatan untuk menggunakan model yang lebih sederhana untuk menyelesaikan tugas kecerdasan buatan [23]. CNN *(Convolutional Neural Network)* juga memiliki kekurangan yaitu tidak dapat bekerja dengan data skuensial yang panjang. Karena CNN tidak memiliki memori, sehingga tidak dapat menyimpan informasi tentang makna kata. Kelemahan ini dapat diatasi dengan menggunakan model LSTM adalah sejenis arsitektur RNN *(Recurrent Neural Network)* yang dirancang untuk "mempertahankan" nilai-nilai yang telah diperoleh sebelumnya untuk periode tertentu keluaran [24]. Berdasarkan perumusan permasalahan yang telah di jelaskan pada paragraf sebelumnya. Penelitian ini bertujuan untuk melakukan klasifikasi opini pelanggan restoran berbasis aspek, berdasarkan ulasan pelanggan restoran pada *TripAdvisor*, berbahasa Indonesia, menggunakan metode *Long Shosrt-Term Memory (LSTM)* dan *Global Max Pooling* untuk klasifikasi sentimen.

## 2. Metode Penelitian

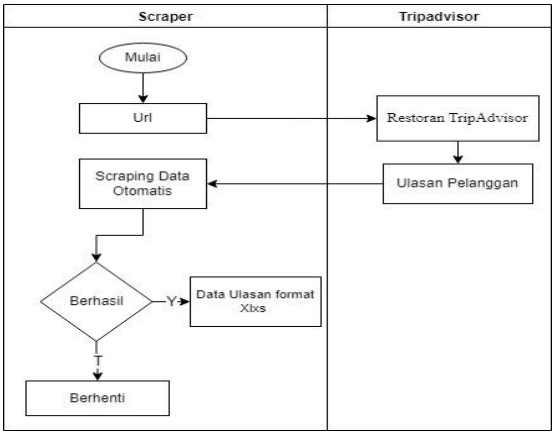
Bab ini membahas metode yang akan digunakan dalam penelitian ini. Metode penelitian terdiri dari beberapa tahap yang dilakukan secara berurutan, yaitu, yaitu *data collection*, *preprocessing*, *word embedding*, *classification*, dan hasil klasifikasi. Gambar 1 menunjukkan langkah-langkah yang diambil dalam penelitian dan memberikan gambaran yang lebih baik tentang prosesnya. Selain itu, metode penelitian akan dibahas secara mendalam di masing-masing sub bab untuk menjelaskan langkah-langkah yang diambil dalam penelitian ini.

### 2.1 Data Collection

Pada penelitian ini, data ulasan pelanggan dikumpulkan melalui metode web *scraping* dari *TripAdvisor*, proses pengambilan *dataset* ini berlangsung dari bulan Juni hingga Juli 2022, dan totalnya sebanyak 1700 ulasan.




Gambar 1. Prosedur Penelitian




Gambar 2. Prosedur Pengambilan Data Penelitian

Data ulasan hanya dikumpulkan dari restoran di Bandar Djakarta Ancol di Indonesia, dan setiap ulasan digunakan dalam Bahasa Indonesia. Untuk melakukan pengumpulan data, *library python* digunakan, yang memudahkan proses *scraping*. Gambar 2 menunjukkan proses pengumpulan data yang digunakan untuk mengumpulkan ulasan pelanggan dari situs web *Tripadvisor*.

Gambar 2 menunjukkan bahwa langkah pertama dalam pengumpulan data adalah mengirimkan permintaan HTTP untuk mengirimkan ulasan data *riview* pelanggan restoran Bandar Djakarta Ancol ke situs *TripAdvisor*. Untuk melakukan ini, modul "permintaan" dalam bahasa pemrograman memungkinkan pengiriman permintaan HTTP ke URL halaman web yang akan *discrape*. *Website Tripadvisor* akan mengirimkan data ulasan dan token yang sesuai dengan permintaan setelah diterima. Setelah data dan token diperoleh sesuai permintaan, data ulasan disimpan dalam format dokumen.xlsx.

## 2.2 Preprocessing

Tahapan text *pre-proccessing* yang merupakan tahapan yang dilalukan untuk membersihkan *dataset* sentimen berupa teks yang terstruktur agar data siap untuk diolah. Selain itu proses ini dapat meningkatkan kinerja dan akurasi model klasifikasi, Dalam sebuah *dataset* ulasan yang berisi kata yang tidak jelas maka kata tersebut akan dihapus[25]. Tahapan pertama *Cleansing* adalah pembersihan atau penghapusan pada semua dokumen yang berisi angka, Url (http://), username (@), tanda pagar (#), delimiter seperti koma (,) dan titik (.) dan juga tanda baca lainya[26].

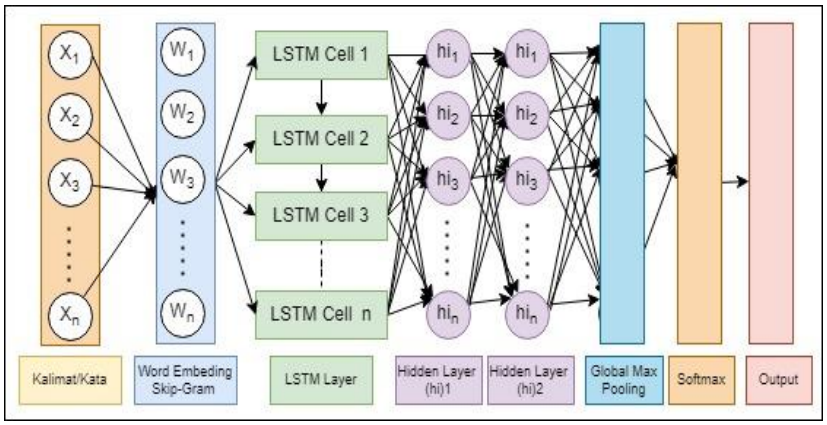
Selanjutnya *Case folding* adalah proses penyeragaman karakter menjadi huruf kecil. Proses ini diperlukan karena kata yang sama dengan font yang berbeda akan dianggap sebagai dua fitur yang berbeda. Hal ini dapat memperbesar dimensi dari token yang dihasilkan namun tidak memberikan arti dari variasi fitur tersebut[27]. Dan *tokenization* adalah proses memisahkan urutan *string* menjadi individu seperti kata, kata kunci, frasa, simbol, dan elemen lain yang dikenal sebagai token. Token dapat berupa kata, frasa, atau bahkan seluruh kalimat[28].

## 2.3 Word Embedding

*Word embedding* ialah istilah yang digunakan untuk representasi kata untuk analisis teks, biasanya dalam bentuk vektor yang bernilai untuk mengkodekan arti kata sedemikian rupa sehingga kata-kata yang lebih dekat dalam ruang vektor diharapkan memiliki makna yang sama[29]. Ada berbagai macam Teknik dalam *word embedding* yang paling populer *Word2Vec (Word to Vector*), metode pembelajaran umum di *Word2Vec*. Model *Skip-gram(SG)* yang ada dalam *Word2vec* digunakan untuk memprediksi kata konteks menggunakan kata informasi yang disebut kata target [30].

## 2.4 Classification

Untuk proses klasifikasi dengan menggunakan pemrograman *Python*. Dengan metode klasifikasi yang diterapkan dalam penelitian ini. Model *Long Short-Term memory (LSTM)* salah satu varian *Recurrent Neural Network (RNN)* yang paling banyak digunakan, dengan menyimpan informasi jangka panjang ke dalam sel memori dan mempelajari informasi kontekstual yang berguna untuk tugas klasifikasi[16]. kinerja RNN tidak memuaskan untuk masalah ketergantungan jangka Panjang, Oleh karena itu, jaringan syaraf tiruan LSTM diusulkan untuk meningkatkan RNN[31]. Pada metode LSTM dengan menambah layer digunakan yaitu *Global Max Pooling* merupakan salah satu teknik pengambilan




Gambar 3. Arsitektur CNN-LSTM

fitur yang sering digunakan dalam jaringan saraf konvolusi(CNN). *Global max pooling* memberikan nilai tunggal dengan mengurangi setiap saluran di peta fitur. *Global max pooling* untuk mengurangi fitur resolusi *output* dan mencegah *overfitting* data [32].

Selain itu metode optimasi yang digunakan adalah *Adam Optimizer* Yang merupakan algoritme pengoptimalan laju pembelajaran adaptif yang dirancang khusus untuk melatih *deep neural network*. Algoritma ini memanfaatkan kekuatan metode kecepatan pembelajaran adaptif untuk menemukan kecepatan pembelajaran individu untuk setiap parameter. Ini menggunakan estimasi momen gradien pertama dan kedua untuk mengadaptasi laju pembelajaran untuk setiap bobot jaringan saraf [33].

Pada penelitian ini proses pelatihan dilakukan dengan beberapa fase, fase propagasi maju, propagasi mundur, dan memperbarui bobot dengan menggunakan metode optimasi *Adam Optimizer*. Dalam proses pelatihan model LSTM dilakukan dengan jumlah n iterasi dan penerapan model LSTM dengan menggunakan *library Tensorflow*. Arsitektur model LSTM dan global max pooling yang dilakukan dalam pelatihan ini dapat dilihat pada Gambar 3.

Proses pelatihan model dilakukan dengan menerapkan beberapa lapisan, yaitu lapisan *Word Embedding* dengan menggunakan *Skip-Gram* kemudian masuk ke metode LSTM selain itu aktivasi yang digunakan adalah aktivasi *ReLu* pada *hidden layer*, lapisan ini menggantikan angka *negative* dari lapisan konvolusi dengan 0 (nol) dan *global max pooling* layer untuk menghindari *overfitting* dan mengurangi beban komputasi, untuk selanjutnya *output* menggunakan *Softmax* yang digunakan dalam mendapatkan hasil klasifikasi. Optimasi yang digunakan dalam metode LSTM yaitu metode *Adam optimizer* untuk mengoptimalkan parameter jaringan saraf.

### 2.5 Evalution dan Testing

Untuk mengevaluasi metode yang diusulkan, dengan menggunakan *Confusion Matrix* merupakan metode untuk mengukur dan menilai tingkat ketepatan sebuah model. Dalam melakukan beberapa eksperimen dengan mengatur setiap model untuk dijalankan sepuluh kali untuk kedua set data. Pada Tabel 1 merupakan *Confusion Matrik* yang digunakan pada penelitian ini [34]. Dengan menggunakan *accuracy (A), precision (P), recall (R), dan $F_1$-score (F1)*. Sebagai metrik evaluasi kinerja seperti pada (1) sampai (4). Di mana *TP* mewakili benar-positif yaitu jumlah sampel positif yang diprediksi dengan benar; *TN* mewakili benar-negatif, jumlah sampel negatif diprediksi dengan benar; *FP* mewakili positif palsu, jumlah sampel negatif yang diprediksi salah menjadi positif; dan *FN* mewakili negatif palsu, jumlah sampel positif yang diprediksi salah menjadi negatif [35].

Untuk menghitung akurasi dari prediksi yang benar dapat menggunakan formula 1. Dengan TP menunjukkan nilai positif benar di tambah TN merupakan nilai sentimen benar di bagi total TP ditambah TN ditambah FP ditambah FN. Presisi digunakan untuk mengukur sampel kelas positif yang diklasifikasi dengan benar dan didefinisikan pada formula 2, dengan TP dan FP menunjukkan jumlah positif benar dan positif palsu. *Recall* pada formula 3 digunakan untuk menghitung semua sampel positif, dengan TP merupakan benar positif dan di bagi jumlah dari TP ditambah FN.

Tabel 1. *Confusion Matrix*

| | | **Kelas Aktual** | |
|---|---|---|---|
| | | Positif | Negatif |
| Kelas Predikasi | Positif | TP (True Positive) | FN (False Negative) |
| | Negatif | FP (False Positive) | TN (True Negative) |

*F-Score* mempertimbangkan dengan menghitung keseimbangan presisi dan daya ingat. Dapat dilihat pada formula 4 adalah rata-rata antara *precision* dan *recall* yang menjelaskan seberapa baik sistem otomatis mampu menyeimbangkan kinerja antara *precision* dan *recall*. Pada dasarnya jika baik *recall* dan *precision* menjadi 1, yang mengarah ke *F-score* sama dengan 1, jika salah satunya 0, *f-score* adalah 0, nilai 1 menunjukkan klasifikasi yang sempurna, yang berarti

Tabel 2. Contoh Data Hasil *Web Scraping*

| No | Komentar |
|---|---|
| 1 | Makanan enak, harga sesuai.. Resto yang harus dikunjungi ketika ke Jakarta bersama rekan dan keluarga.. |
| 2 | salah satu tempat yang nyaman untuk dikunjungi ketika pergi ke jakarta . oke untuk berkumpul dengan keluarga |
| 3 | Makanannya enak enak, harga sebanding dengan rasa. Tempat nyaman. Hanya kalau pas weekend ke sini sangat ramai. Over all good |
| 4 | Tempat yang sangat tepat untuk berkumpul dengan keluarga... terlebih setelah seharian explore taman rekreasi Ancol.. tempatnya sangat luas nyaman dan untuk kebersihannya sangat terjaga untuk restaurant yang kapasitas besar dan high turned over customer yg akan dine in. Pelayanan diberikan sangat friendly dan staff memberikan... |
| 5 | untuk pelayanannya sangat bagus, cepat dan ramah. dan untuk makanannya enak enak apalagi untuk pemandangannya yang bagus untuk foto foto. |
| 6 | Suasananya asikk bgt sangat worth it makan malam disini . Makananan enak dan cocok sekali untuk keluarga ataupun teman |

semua prediksi benar, dan nilai 0 menunjukkan klasifikasi yang sangat buruk, yang berarti tidak ada prediksi yang benar.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F_1 - score = 2\frac{PxR}{PxR} \quad (4)$$

## 3. Hasil dan Pembahasan

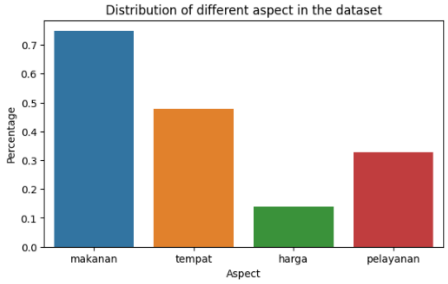
Model *Long Short-Term Memory* diterapkan pada penelitian ini dalam kasus analisis sentimen berbasis aspek ulasan konsumen restoran pada *TripAdvisor* dengan aktivasi *ReLu* pada *hidden layer*, *Global max pooling layer* untuk mengurangi beban komputasi, dan *Adam optimizer* ditranslasikan dalam Bahasa pemrograman *python* yang menghasilkan sistem informasi berbasis web. Sistem informasi tersebut menampilkan hasil klasifikasi sentimen ulasan pelanggan dengan polarisasi negatif, setengah negatif, netral, setengah positif, dan positif.

### 3.1. Pengumpulan Data

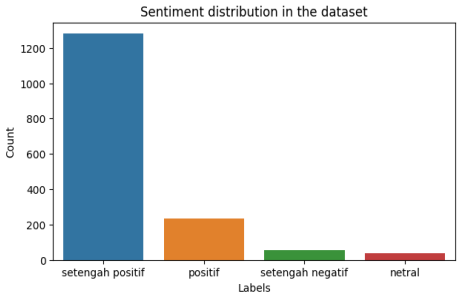
Data dalam penelitian ini, dikumpulkan melalui *scraping* web *TripAdvisor* dari Juni hingga Juli 2022. Jumlah total data ulasan adalah 1700. Pengumpulan data ini dilakukan menggunakan bahasa pemrograman *Phyton*. Tabel 2 menunjukkan contoh data yang dihasilkan dari web *scraping*, yang menunjukkan atribut yang terkait dengan ulasan pelanggan. Data yang dihasilkan dari web *scraping* ini akan digunakan sebagai dasar untuk analisis sentimen berbasis aspek yang dilakukan dalam penelitian ini. Penelitian ini berhasil mengumpulkan data ulasan *TripAdvisor* cukup besar dengan menggunakan metode web scrapping/

```
0       [makanan, enak, harga, sesuai, resto, yang, ha...
1       [salah, satu, tempat, yang, nyaman, untuk, dik...
2       [makanannya, enak, enak, harga, sebanding, den...
3       [tempat, yang, sangat, tepat, untuk, berkumpul...
4       [untuk, pelayanannya, sangat, bagus, cepat, da...
                              ...
1735    [setelah, mengunjungi, pantai, ancol, anda, se...
1736    [makanan, di, bandar, djakartar, adalah, pasti...
1737    [sangat, menarik, tempat, untuk, makan, dan, b...
1738    [saya, suka, makanan, laut, dan, bandar, jakar...
1739    [saya, tahu, restoran, ini, karena, kakak, say...
```

Gambar 4. Contoh Hasil *Preprocessing* Data




Gambar 5. Datset Aspek



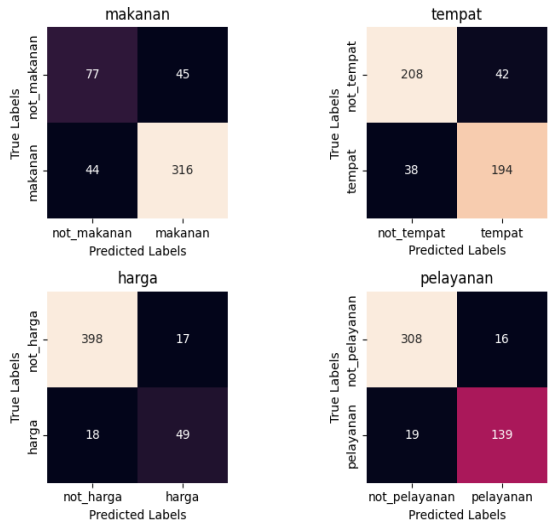

Gambar 6. Dataset Sentimen

Ini memberikan dasar yang kuat untuk melakukan analisis sentimen dan menemukan elemen yang mempengaruhi ulasan pelanggan.

### 3.2. Preprocessing Data

Pembersihan data dilakukan dengan menghapus beberapa ulasan duplikat, simbol, tag HTML, tanda baca, spasi, angka, karakter khusus, disebut dengan *Cleansing.* Selanjutnya proses *case folding* untuk mengubah semua huruf menjadi huruf kecil semua agar kata yang sama akan tetapi penulisan hurufnya berbeda tidak menjadi data ganda. Kemudian untuk memecah suatu kalimat, paragraf atau dokumen menjadi beberapa potongan yang disebut dengan *token* atau dikenal dengan nama *Tokenization.* Proses *preprocessing* menghasilkan data ulasan pelanggan bersih yang selanjutnya dijadikan sebagai data latih dan data uji. Hasil dari proses *preprocessing* data pada penelitian ini dapat dilihat pada Gambar 4.

*Preprocessing* selesai, tahap selanjutnya menyiapkan data latih dan data uji kemudian proses *word embedding* dengan menggunakan *word2vec* yang digunakan *skip-gram* untuk menghitung manual dengan menggunakan *python* pembobotan dalam data ulasan.




Gambar 7. *Confusion Matrix* Aspek

Setelah proses *preprocessing* kemudian dilakukan proses pembuangan kalimat yang tidak memiliki aspek sentimen, data bersih menghasilkan 1605 data, pada Gambar 5 *dataset* aspek, dan Gambar 6 *dataset* sentimen, merupakan hasil pembagian data berdasarkan aspek dan sentimennya. dengan kategori sentimen jumlah data pada netral 36 data, untuk positif 235 data, setengah negatif 54 data, setengah positif 1280 data.

### 3.3. Evalution dan Testing

Pada tahap evaluasi, hasil akurasi digunakan untuk mengidentifikasi nilai klasifikasi yang berhasil dibangun pada saat prosedur model klasifikasi sebelumnya. Pendekatan validasi 10 *K-fold* adalah yang digunakan untuk evaluasi. Untuk memastikan bahwa hasilnya tepat, diperlukan pengukuran evaluasi dalam penelitian ini *confusion matrix* digunakan untuk evaluasi sistem. Hasil evaluasi sistem menggunakan 1605 sampel data, yang ditampilkan terdiri dari dua *confusion matrix* yaitu *confusion matrix* sentimen dan *confusion matrix* aspek. Hasil *confusion matrix* aspek kategori yang digunakan untuk menghasilkan akurasi, *precision*, *f1-score* dan *recall*, dapat dilihat pada Gambar 7.

Pada penelitian ini, *confusion matrix* digunakan untuk evaluasi sistem. Hasil eksperimen kategori aspek makanan untuk hasil *confusion matrix* dapat dilihat pada Tabel 3. Berdasarkan jumlah data yang diperoleh, terdapat 77 review yang salah prediksi dengan nilai negatif (*true negative*) dan 316 prediksi benar dengan nilai positif (*true positive*). Selain itu, terdapat 45 review yang salah prediksi dengan nilai negatif (*false negative*) serta 44 ulasan memprediksi salah dengan nilai positif (*false positive*).

Untuk selanjutnya, perhitungan *confusion matrix* pada aspek kategori tempat telah dilakukan. Hasil perhitungan *confusion matrix* dapat dilihat pada Tabel 4. Berdasarkan hasil terdapat 208 hasil prediksi yang salah dengan nilai negatif (*true negative*), serta 194 prediksi yang benar dengan nilai positif (*true positive*). Selain itu, terdapat 42 prediksi yang salah dengan nilai negatif (*false negative*), dan terakhir, terdapat 38 prediksi yang salah diprediksi dengan nilai positif (*false positive*).

Tabel 3. Hasil Confusion Matrix Aspek Makanan

| Aspek | Precision | recall | F1-score |
|---|---|---|---|
| Makanan | 88% | 88% | 88% |

Tabel 4. Hasil Confusion Matrix Aspek Tempat

| Aspek | Precision | recall | F1-score |
|---|---|---|---|
| Tempat | 84% | 82% | 83% |

Tabel 5. Hasil *Confusion Matrix* Aspek Harga

| Aspek | Precision | recall | F1-score |
|---|---|---|---|
| Harga | 73% | 74% | 74% |

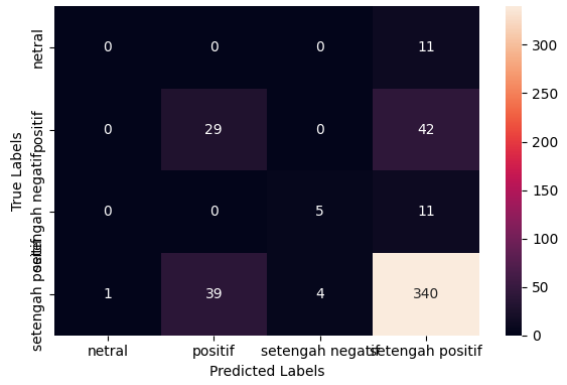
Tabel 6. Hasil *Confusion Matrix* Aspek Pelayanan

| Aspek | Precision | recall | F1-score |
|---|---|---|---|
| Pelayanan | 88% | 90% | 89% |

Untuk perhitungan selanjutnya, pada aspek kategori harga, berdasarkan perhitungan dari *confusion matrix* dapat dilihat pada Tabel 5. Telah dilakukan perhitungan *confusion matrix* dengan jumlah data sebanyak 398 review yang diprediksi salah dengan nilai negatif (*true negative*) selanjutnya dengan jumlah data 49 review yang diprediksi benar dengan nilai positif (*true positive*). Selain itu, terdapat 17 review diprediksi salah dengan nilai negatif (*false negative*) serta jumlah data 18 ulasan diprediksi salah dengan nilai positif (*false positive*).

Pada perhitungan selanjutnya untuk aspek kategori pelayanan, hasil perhitungan *confusion matrix* dapat dilihat pada Tabel 6. Terdapat jumlah data dengan 308 ulasan yang diprediksi salah dengan nilai negatif (*true negative*), selanjutnya dengan jumlah data 139 ulasan yang diprediksi benar dengan nilai positif (*true positive*). Namun terdapat 16 ulasan yang diprediksi salah dengan nilai negatif (*false negative*), yang berarti ada sejumlah ulasan dengan kategori pelayanan yang seharusnya diprediksi positif, namun diprediksi sebagai negatif oleh model. Selain itu, terdapat 19 ulasan yang diprediksi salah dengan nilai positif (*false positive*), yang berarti ada sejumlah ulasan dengan kategori pelayanan yang seharusnya prediksi negatif, namun diprediksi sebagai positif oleh model.

Berdasarkan hasil perhitungan evaluasi dengan menggunakan metode *Confusion Matrix*, dapat dilihat pada Tabel 7 bahwa penggunaan model LSTM dan *global max pooling* melakukan analisis sentimen berbasis aspek ulasan pelanggan restoran pada *TripAdvisor*. Analisis yang dilakukan dalam sentimen pada pelanggan secara spesifik terhadap aspek kategori




Gambar 8. *Confusion Matrix* Sentimen

Tabel 7. Hasil *Confusion Matrix* Aspek

| | Accuracy | Precision | F1-Score | Recall |
|---|---|---|---|---|
| ABSA-LSTM | 78.7% | 85% | 85% | 85% |

Tabel 8. Hasil *Confusion Matrix* Sentimen

| | Accuracy | Precision | F1-Score | Recall |
|---|---|---|---|---|
| Sentiment | 78% | 81% | 79% | 78% |

makanan, tempat, harga, dan layanan sebagai variabel penilaian. Hasil perhitungan *confusion matrix* pada aspek kategori menunjukkan bahwa akurasi yang diperoleh mencapai 78.7%. Artinya, model yang digunakan berhasil mengklasifikasikan aspek kategori yang telah ditentukan dengan tingkat akurasi yang signifikan.

Selain itu, evaluasi menggunakan *confusion matrix* untuk sentimen tidak jauh berbeda dengan aspek kategori. Dalam evaluasi sentimen, setiap kelas prediksi dan kelas aktual akan dibandingkan dengan nilai variabel di kelas lain. Hasil evaluasi ini terlihat dalam *confusion matrix* yang dapat ditemukan pada Gambar 8. *Confusion matrix* untuk sentimen menunjukkan bahwa terdapat 4 ulasan yang diprediksi dengan benar (TP) sebagai sentimen negatif, serta 278 ulasan yang diprediksi benar (TP) sebagai sentimen positif dihentikan. Selain itu, terdapat 4 ulasan yang diprediksi dengan benar (TP) sebagai sentimen netral. Namun, tidak ada ulasan 0 yang diprediksi dengan benar (TP)sebagai sentiment positif.

Hasil perhitungan evaluasi menggunakan metode *Confusion Matrix* pada model LSTM untuk analisis sentimen berbasis aspek memberikan informasi yang penting. Evaluasi ini dilakukan dengan menggunakan kelas sentimen negatif, setengah negatif, netral, setengah positif, dan positif. Tabel 8 menunjukkan hasil perhitungan evaluasi ini, yang memberikan gambaran yang lebih rinci tentang kinerja model dalam mengklasifikasikan sentimen pada setiap kategori. Dengan melihat hasil ini, dapat diambil Langkah-langkah perbaikan yang diperlukan untuk meningkatkan akurasi dan keandalan model dalam mengidentifikasi sentimen pada analisis berbasis aspek.

## 4. Kesimpulan

Hasil penelitian ini menunjukkan bahwa analisis sentimen berbasis aspek untuk ulasan pelanggan sangat penting untuk mengetahui klasifikasi opini pelanggan restoran di *TripAdvisor*. Polaritas sentimen digunakan untuk mewakili sentimen yang terkandung dalam ulasan, dan aspek kategori berfungsi untuk mengidentifikasi atribut topik dari ulasan. Eksperimen analisis sentimen menunjukkan hasil dengan akurasi 78.7% dan analisis sentimen berbasis aspek dengan nilai akurasi 78%. Model LSTM dan *global max pooling* menunjukkan hasil kategori dan polaritas sentimen yang merepresentasikan perasaan pelanggan. Penggunaan metode LSTM dan *global max pooling* dalam proses klasifikasi menghasilkan informasi terkait opini pelanggan restoran berdasarkan empat aspek yang ditentukan yaitu makanan, pelayanan, tempat, harga, dan polaritas sentimen yang digunakan yaitu negatif, setengah negatif, netral, setengah positif, dan positif.

Kategori aspek dapat membantu mengidentifikasi aspek mana yang harus dipertahankan dengan mempertimbangkan sentimen positif dan aspek mana yang harus ditingkatkan dengan mempertimbangkan sentimen negatif. Hal ini menunjukkan bahwa polaritas sentimen dan aspek review pelanggan merupakan hal penting yang saling berhubungan, karena mengetahui sentimen saja tanpa mengetahui aspek review tidak cukup untuk mengetahui pendapat pelanggan restoran. Adapun yang dapat dilakukan untuk penelitian selanjutnya terkait analisis sentimen berbasis aspek, pada penelitian selanjutnya dapat menggunakan lebih dari satu ulasan restoran. Karena penelitian ini terbatas pada satu review restoran saja.

## Daftar Pustaka

[1] R. Filieri, F. Acikgoz, V. Ndou, dan Y. Dwivedi, "Is TripAdvisor still relevant? The influence of review credibility, review usefulness, and ease of use on consumers' continuance intention," *Int. J. Contemp. Hosp. Manag.*, vol. 33, no. 1, hal. 199–223, 2021, doi: 10.1108/IJCHM-05-2020-0402.

[2] H. Li, R. Qi, H. Liu, F. Meng, dan Z. Zhang, "Can time soften your opinion? The influence of consumer experience valence and review device type on restaurant evaluation," *Int. J. Hosp. Manag.*, vol. 92, no. 1, hal. 0278–4319, 2021, doi: 10.1016/j.ijhm.2020.102729.

[3] Statista, "Tripadvisor: number of reviews 2021 | Statista," *Tripadvisor*, 2022. [Daring]. Tersedia pada: https://www.statista.com/statistics/684862/tripadvisor-number-of-reviews/. [Diakses: 24-Sep-2022].

[4] D. Liang, Z. Dai, M. Wang, dan J. Li, "Web celebrity shop assessment and improvement based on online review with probabilistic linguistic term sets by using sentiment analysis and fuzzy cognitive map," *Fuzzy Optim. Decis. Mak.*, vol.

19, no. 4, hal. 561–586, 2020, doi: 10.1007/s10700-020-09327-8.

[5] J. Ara, M. T. Hasan, A. Al Omar, dan H. Bhuiyan, "Understanding Customer Sentiment: Lexical Analysis of Restaurant Reviews," *2020 IEEE Reg. 10 Symp. TENSYMP 2020*, vol. 10, no. 1, hal. 295–299, 2020, doi: 10.1109/TENSYMP50017.2020.9230712.

[6] S. Aktas Polat dan S. Polat, "Discovery of factors affecting tourists' fine dining experiences at five-star hotel restaurants in Istanbul," *Br. Food J.*, vol. 124, no. 1, hal. 221–238, 2021, doi: 10.1108/BFJ-02-2021-0138.

[7] M. Gebbels, A. McIntosh, dan T. Harkison, "Fine-dining in prisons: Online TripAdvisor reviews of The Clink training restaurants," *Int. J. Hosp. Manag.*, vol. 95, no. 1, hal. 0278–4319, 2021, doi: 10.1016/j.ijhm.2021.102937.

[8] R. Khorsand, M. Rafiee, dan V. Kayvanfar, "Insights into TripAdvisor's online reviews: The case of Tehran's hotels," *Tour. Manag. Perspect.*, vol. 34, no. 1, hal. 2211–9736, 2020, doi: 10.1016/j.tmp.2020.100673.

[9] P. Abeysinghe dan T. Bandara, "A novel self-learning approach to overcome incompatibility on TripAdvisor reviews," *Data Sci. Manag.*, vol. 5, no. 1, hal. 1–10, Mar 2022, doi: 10.1016/J.DSM.2022.02.001.

[10] A. Valdivia *dkk.*, "Inconsistencies on TripAdvisor reviews: A unified index between users and Sentiment Analysis Methods," *Neurocomputing*, vol. 353, no. 1, hal. 3–16, Agu 2019, doi: 10.1016/J.NEUCOM.2018.09.096.

[11] O. Sharif, M. M. Hoque, dan E. Hossain, "Sentiment Analysis of Bengali Texts on Online Restaurant Reviews Using Multinomial Naïve Bayes," *1st Int. Conf. Adv. Sci. Eng. Robot. Technol. 2019, ICASERT 2019*, vol. 2019, no. Icasert, hal. 1–6, 2019, doi: 10.1109/ICASERT.2019.8934655.

[12] J. A. Kumar, T. E. Trueman, dan E. Cambria, "Gender-based multi-aspect sentiment detection using multilabel learning," *Inf. Sci. (Ny).*, vol. 606, no. 1, hal. 453–468, 2022, doi: 10.1016/j.ins.2022.05.057.

[13] C. Yang, H. Zhang, B. Jiang, dan K. Li, "Aspect-based sentiment analysis with alternating coattention networks," *Inf. Process. Manag.*, vol. 56, no. 3, hal. 463–478, 2019, doi: 10.1016/j.ipm.2018.12.004.

[14] D. H. Pham dan A. C. Le, "Exploiting multiple word embeddings and one-hot character vectors for aspect-based sentiment analysis," *Int. J. Approx. Reason.*, vol. 103, hal. 1–10, Des 2018, doi: 10.1016/J.IJAR.2018.08.003.

[15] J. Joseph, S. Vineetha, dan N. V. Sobhana, "A survey on deep learning based sentiment analysis," *Mater. Today Proc.*, vol. 58, no. 1, hal. 456–460, 2022, doi: 10.1016/j.matpr.2022.02.483.

[16] M. H. Su, C. H. Wu, K. Y. Huang, dan T. H. Yang, "Cell-Coupled Long Short-Term Memory with L-Skip Fusion Mechanism for Mood Disorder Detection through Elicited Audiovisual Features," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 31, no. 1, hal. 124–135, 2020, doi: 10.1109/TNNLS.2019.2899884.

[17] L. Li, T. T. Goh, dan D. Jin, "How textual quality of online reviews affect classification performance: a case of deep learning sentiment analysis," *Neural Comput. Appl.*, vol. 32, no. 9, hal. 4387–4415, Mei 2020, doi: 10.1007/S00521-018-3865-7/TABLES/10.

[18] A. Krishna, V. Akhilesh, A. Aich, dan C. Hegde, "Sentiment Analysis of Restaurant Reviews Using Machine Learning Techniques," *Lect. Notes Electr. Eng.*, vol. 545, no. 1, hal. 687–696, 2019, doi: 10.1007/978-981-13-5802-9_60.

[19] N. Hossain, M. R. Bhuiyan, Z. N. Tumpa, dan S. A. Hossain, "Sentiment Analysis of Restaurant Reviews using Combined CNN-LSTM," *2020 11th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2020*, vol. 10, hal. 1–5, 2020, doi: 10.1109/ICCCNT49239.2020.9225328.

[20] L. Li, L. Yang, dan Y. Zeng, "Improving Sentiment Classification of Restaurant Reviews with Attention-Based Bi-GRU Neural Network," *Symmetry (Basel).*, vol. 13, no. 8, hal. 1517, 2021.

[21] H. H. Do, P. W. C. Prasad, A. Maag, dan A. Alsadoon, "Deep Learning for Aspect-Based Sentiment Analysis: A Comparative Review," *Expert Syst. Appl.*, vol. 118, hal. 272–299, 2019, doi: 10.1016/j.eswa.2018.10.003.

[22] J. Sheela dan B. Janet, "An abstractive summary generation system for customer reviews and news article using deep learning," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 7, hal. 7363–7373, 2021, doi: 10.1007/s12652-020-02412-1.

[23] S. Sohangir, D. Wang, A. Pomeranets, dan T. M. Khoshgoftaar, "Big Data: Deep Learning for financial sentiment analysis," *J. Big Data*, vol. 5, no. 3, hal. 2–25, 2018, doi: 10.1186/s40537-017-0111-6.

[24] P. F. Muhammad, R. Kusumaningrum, dan A. Wibowo, "Sentiment Analysis Using Word2vec and Long Short-Term Memory (LSTM) for Indonesian Hotel Reviews," *Procedia Comput. Sci.*, vol. 179, no. 2020, hal. 728–735, 2021, doi: 10.1016/j.procs.2021.01.061.

[25] S. Bessou dan R. Aberkane, "Subjective Sentiment Analysis for Arabic Newswire Comments," *J. Digit. Inf. Manag.*, vol. 17, no. 5, hal. 289–295, 2019, doi: 10.6025/jdim/2019/17/5/289-295.

[26] M. Shaheen, "Sentiment Analysis on Mobile Phone Reviews Using Supervised Learning Techniques," *Int. J. Mod. Educ. Comput. Sci.*, vol. 11, no. 7, hal. 32–43, 2019, doi: 10.5815/ijmecs.2019.07.04.

[27] S. Khomsah, "Naive Bayes Classifier Optimization on Sentiment Analysis of Hotel Reviews," *J. Penelit. Pos dan Inform.*, vol. 10, no. 2, hal. 157–168, 2020, doi: 10.17933/jppi.2020.100206.

[28] T. Ul Haque, N. Nawal Saber, dan F. Muhammad Shah, "Sentiment Analysis on Large Scale Amazon Product Reviews," in *2018 IEEE International Conference on Innovative Research and Development (ICIRD)*, 2018, no. May, hal. 1–6.

[29] N. Kanakaris, N. Giarelis, I. Siachos, dan N. Karacapilidis, "Making personnel selection smarter through word embeddings: A graph-based approach," *Mach. Learn. with Appl.*, vol. 21, no. 9, hal. 100–214, 2021, doi: 10.1016/j.mlwa.2021.100214.

[30] S. Amin, M. Irfan Uddin, M. Ali Zeb, A. A. Alarood, M. Mahmoud, dan M. H. Alkinani, "Detecting dengue/flu infections based on tweets using LSTM and word embedding," *IEEE Access*, vol. 8, no. 1, hal. 189054–189068, 2020, doi: 10.1109/ACCESS.2020.3031174.

[31] X. Song *dkk.*, "Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model," *J. Pet. Sci. Eng.*, vol. 186, no. 1, hal. 0920–4105, 2020, doi: 10.1016/j.petrol.2019.106682.

[32] K. Ullah, A. Rashad, M. Khan, Y. Ghadi, H. Aljuaid, dan Z. Nawaz, "A Deep Neural Network-Based Approach for Sentiment Analysis of Movie Reviews," *Complexity*, vol. 2022, 2022, doi: 10.1155/2022/5217491.

[33] S. M. Qaisar, "Sentiment Analysis of IMDb Movie Reviews Using Long Short-Term Memory," *2020 2nd Int. Conf. Comput. Inf. Sci. ICCIS 2020*, hal. 12–15, 2020, doi: 10.1109/ICCIS49240.2020.9257657.

[34] K. Goseva-Popstojanova dan J. Tyo, "Identification of security related bug reports via text mining using supervised and unsupervised classification," *Proc. - 2018 IEEE 18th Int. Conf. Softw. Qual. Reliab. Secur. QRS 2018*, hal. 344–355, 2018, doi: 10.1109/QRS.2018.00047.

[35] S. Chotirat dan P. Meesad, "Part-of-Speech tagging enhancement to natural language processing for Thai wh-question classification with deep learning," *Heliyon*, vol. 7, no. 10, hal. 2405–8440, Okt 2021, doi: 10.1016/J.HELIYON.2021.E08216.

*Halaman ini sengaja dikosongkan*