



Kuantifikasi Prioritas Pengembangan Fitur Aplikasi dengan Kerangka *Outcome-Driven Innovation*: Studi Kasus Aplikasi Pembelajaran Bahasa Asing

Kevin Bastian Sirait^{1*}, Nicholas Dickson², Romindo³

^{1*,2,3}Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Pelita Harapan, Medan, Indonesia

Email: ^{1,*}kevin.sirait@uph.edu, ²03081210020@student.uph.edu, ³romindo@uph.edu

Abstract

This study aims to identify user satisfaction, importance, and development priorities for the English Language Speech Assistant (ELSA) application by integrating Outcome-Driven Innovation (ODI) and sentiment analysis methods. Using reviews collected from the Google Play Store via ScrapeStorm, two sentiment models were applied. Namely, Valance Aware Dictionary and sEntiment Reasoner (VADER) for overall sentiment polarity and Aspect-Based Sentiment Analysis (ABSA) for feature-specific sentiment extraction. The result show that 86.23% of user reviews express positive sentiments, reflecting high satisfaction with ELSA's functionality and learning experience. Correlation testing confirms strong validity between VADER and ABSA models ($r = 0.98$), and moderate alignment with user rating ($r \approx 0.55$), confirming both construct and external validity. The ODI analysis further identifies Pronunciation as the highest-priority aspect (opportunity = 12.1), followed by Learning (opportunity = 2.6), highlighting key areas for innovation. Other aspects, such as Support and Subscription, show balanced performance with low opportunity values. The integration of ODI and sentiment analysis provide a data-driven framework for product improvement, enabling developers to prioritize enhancements based on user-perceived value. The findings contribute to informed strategic decision-making in the development of digital language-learning applications.

Keywords: Mobile Application, Outcome-Driven Innovation, User-Centered Design, Features Development.

Abstrak

Penelitian ini bertujuan untuk mengidentifikasi tingkat kepuasan, tingkat kepentingan, serta prioritas pengembangan pengguna aplikasi *English Language Speech Assistant* (ELSA) melalui integrasi kerangka *Outcome-Driven Innovation* (ODI) dan analisis sentimen. Data penelitian diperoleh dari ulasan pengguna di Google Play Store menggunakan ScrapeStorm, kemudian dianalisis dengan dua model sentimen, yaitu *Valence Aware Dictionary and sEntiment Reasoner* (VADER) untuk mengukur polaritas sentimen secara keseluruhan dan *Aspect-Based Sentiment Analysis* (ABSA) untuk menilai sentimen berdasarkan aspek fitur. Hasil penelitian menunjukkan bahwa 86.23% ulasan pengguna berupa positif, mencerminkan tingkat kepuasan yang tinggi terhadap fungsi dan pengalaman belajar dalam aplikasi ELSA. Uji korelasi menunjukkan validitas yang kuat antara model VADER dan ABSA ($r = 0.98$) serta hubungan moderat dengan penilaian pengguna ($r \approx 0.55$), yang menegaskan validitas konstruk dan eksternal. Analisis ODI mengidentifikasi *Pronunciation* sebagai aspek dengan prioritas tertinggi (opportunity = 12.1), diikuti oleh *Learning* (opportunity = 2.6) sebagai area pengembangan utama. Aspek lain seperti *Support* dan *Subscription* menunjukkan performa seimbang dengan peluang pengembangan rendah. Integrasi ODI dan analisis sentimen menyediakan kerangka berbasis data untuk perbaikan produk, membantu pengembang memprioritaskan inovasi berdasarkan nilai yang dirasakan pengguna.

Kata Kunci: Aplikasi Mobile, Outcome-Driven Innovation, User-Centered Design, Pengembangan Fitur.

1. PENDAHULUAN

Perkembangan terhadap performa, kapabilitas, dan fungsionalitas aplikasi telah memberikan dampak signifikan dalam kehidupan sehari-hari, dimana hal tersebut didorong dengan adanya umpan balik dan interaksi pengguna atas kelebihan dan kekurangan suatu aplikasi (Fu et al., 2023). Interaksi tersebut membangun fondasi terhadap proses pengembangan yang mengacu kepada peningkatan kepuasan dan loyalitas pengguna (Ribeiro et al., 2019). Pendekatan yang berorientasi terhadap pengguna (i.e., *user-centered design*) berfokus dalam memastikan suatu produk tetap selaras dengan ekspektasi para pengguna, dengan tujuan untuk mempertahankan dan meningkatkan retensi pengguna dalam jangka panjang.

Berdasarkan hal tersebut, *Outcome-Driven Innovation* (ODI) merupakan kerangka yang menekankan tahapan pengembangan yang dituntun dan mengacu kepada kebutuhan dan ekspektasi pengguna (Serna-Herrera et al., 2025). Pendekatan ini relevan didalam konteks dimana integrasi teknologi dan kompetisi yang terdapat didalam pasar aplikasi menuntut adanya perkembangan dan adaptasi yang berkelanjutan (Martens & Maalej, 2019). Ribeiro et al. (2019) menjelaskan bahwa kerangka ODI, yang menerapkan konsep *Jobs-To-Be-Done* (JTBD), memberikan kemampuan kepada organisasi dalam mengidentifikasi dan memprioritaskan produk yang mempunyai dampak langsung dan memenuhi ekspektasi para pengguna. Serta, entitas bisnis dapat menghasilkan inovasi yang meningkatkan kepuasan dan loyalitas pengguna dengan cara menghubungkan ekspektasi pengguna dengan target pengembangan yang terukur secara sistematis.

Secara esensi, kerangka ODI mengikuti filosofi *user-centered design* yang bergantung kepada umpan balik para pengguna, respons pasar, dan keselarasan secara berkelanjutan dengan kebutuhan pengguna (Haug et al., 2023). Penerapan kerangka ODI memberikan kapabilitas kepada organisasi dalam menemukan area pengembangan yang dapat memberikan nilai tambah dari peningkatan performa layanan dan pengalaman penggunaan (i.e., *user experience*). Sebagai contoh, Handayani et al. (2022) menemukan bahwa integrasi ODI dalam produk pembelajaran tidak hanya meningkatkan utilitas dari produk tersebut, melainkan efektivitas pembelajaran pengguna. Selaras dengan hal tersebut, Valtonen et al. (2021) menunjukkan bahwa aplikasi edukasi berbasis ODI memberikan fleksibilitas dan, disaat yang sama, menjaga relevansi fitur aplikasi terhadap perkembangan atau perubahan pembelajaran pelajar.

Atas kelebihan tersebut, penelitian terdahulu (seperti Awajan et al., 2021; Corral Abad et al., 2021; Nam et al., 2021) telah melakukan integrasi ODI dengan analisa sentimen untuk mengubah ulasan para pengguna menjadi target pengembangan yang terukur. Nam et al. (2021), dalam konteks perhotelan, menekankan bahwa integrasi tersebut memberikan dasar ukuran dalam memprioritaskan pengembangan suatu produk berdasarkan ketidakpuasan yang terkandung didalam ulasan para pengguna. Selaras dengan temuan tersebut, Fu et al. (2023) dan Awajan et al. (2021) telah mengidentifikasi fitur utama yang harus menjadi prioritas pengembangan dalam rangka meningkatkan loyalitas dan *engagement* para pengguna; pola yang sama juga ditemukan pada penelitian yang dilakukan oleh Corral Abad et al. (2021) didalam konteks edukasi. Secara esensi, penelitian terdahulu telah memberikan argumentasi atas dasar bahwa integrasi antara ODI dan analisa sentimen secara efektif menentukan prioritas pengembangan berdasarkan sentimen, kebutuhan, dan ekspektasi pengguna.

Walaupun terdapat dampak positif atas penerapan kerangka ODI dengan analisa sentimen, terdapat kesenjangan dari sisi integrasi secara sistematis. Seperti yang telah dijelaskan oleh Nam et al. (2021), penerapan analisa sentimen tradisional berfokus kepada identifikasi sentimen atas opini para pengguna namun kurang memiliki

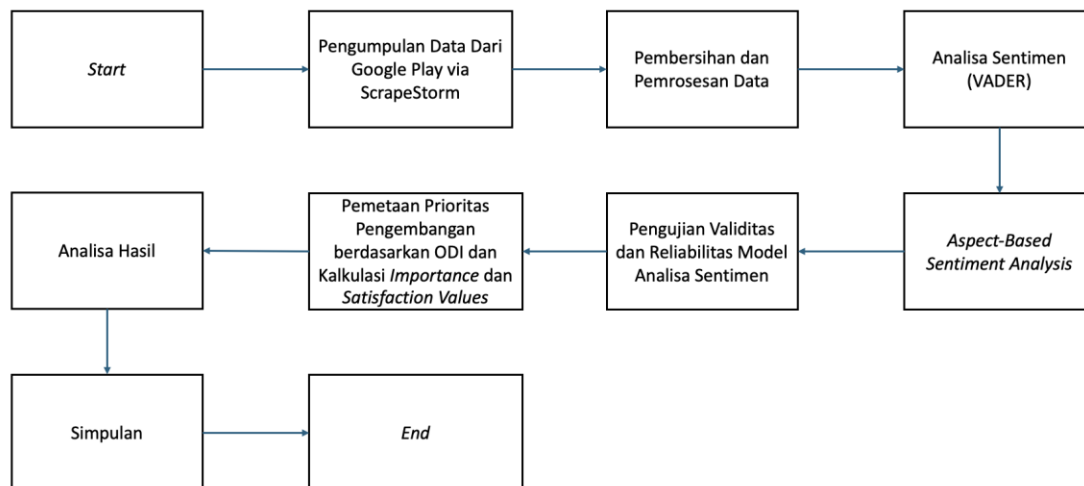
keterhubungan dengan kerangka inovasi seperti ODI, hal ini menyebabkan tidak lengkapnya informasi atas bagaimana suatu sentimen ter-asosiasi dengan suatu fitur spesifik. Pendekatan konvensional secara *lexicon* atau *rule-based* memiliki kelemahan dalam memahami kompleksitas yang terkandung didalam data tekstual dalam ukuran besar (Awajan et al., 2021). Bukti teoritis menunjukkan bahwa dengan mengkombinasikan pendekatan NLP dengan ODI dapat menghasilkan pemahaman yang lebih mendalam terhadap aspek (atau fitur) yang dapat mempengaruhi tingkat kepuasan dan retensi pengguna (Awajan et al., 2021; Fu et al., 2023). Atas kesenjangan tersebut, penelitian ini bertujuan untuk (1) mengidentifikasi sentimen paling dominan terhadap suatu aplikasi, (2) menentukan aspek utama yang merupakan prioritas berdasarkan kerangka ODI, dan (3) menemukan aspek yang mempunyai dampak paling signifikan atas kepuasan pengguna apabila dikembangkan. Dalam merealisasikan tiga tujuan tersebut, penelitian ini menggunakan data ulasan dari aplikasi *English Language Speech Assistant* (ELSA), aplikasi pembelajaran bahasa asing yang memiliki keterikatan dalam efektivitas pembelajaran, dimana aplikasi tersebut merupakan objek ideal dalam mempelajari bagaimana prioritas pengembangan aplikasi dituntun secara nilai sentimen pengguna (Handayani et al., 2022; Kim, 2021; Liu, 2023; Valtonen et al., 2021)

Penelitian ini berkontribusi terhadap perspektif pengembangan aplikasi dengan cara mengintegrasikan model analisa sentimen dengan kerangka ODI dalam penentuan prioritas pengembangan aplikasi. Secara praktis, penelitian ini memberikan landasan pengembangan kepada para pengembang untuk mengidentifikasi dan memenuhi ekspektasi pengguna dan, disaat yang sama, menghasilkan keputusan strategis dalam disain dan perancangan produk (Kim, 2021; Nam et al., 2021).

2. METODOLOGI PENELITIAN

2.1 Disain Penelitian

Penelitian ini menerapkan pendekatan *mixed-method design* yang menggabungkan interpretasi kualitatif dengan kerangka ODI dan interpretasi kuantitatif analisa sentimen dengan menggunakan model *Valence Aware Dictionary and sEntiment Reasoner* (VADER) dan *Aspect-Based Sentiment Analysis* (ABSA). Analisa yang dilakukan mengikuti pola *sequential explanatory design*, dimana hasil komputasi yang telah diperoleh dari analisa sentimen diinterpretasi dengan menggunakan kerangka ODI dalam rangka untuk mengkontekstualisasi hasil kuantitatif dengan perspektif kualitatif. Secara gambaran besar, disain penelitian yang digunakan pada penelitian ini diilustrasikan pada Gambar 1.



Gambar 1. Alur Tahapan Penelitian

2.2 Sumber dan Pemrosesan Data

Data ulasan para pengguna aplikasi ELSA diperoleh dari Google Play Store dengan menggunakan ScrapeStorm—aplikasi *web-scraping*—dengan jumlah data yang terkumpul sebanyak 6420 data. Data yang telah dikumpulkan meliputi rating, tanggal ulasan, dan ulasan dari para pengguna; yang dimana data struktur dari data tersebut digunakan sebagai fondasi untuk analisa sentimen dan ODI. Pada penelitian ini, hanya ulasan dengan bahasa Inggris yang digunakan dalam rangka untuk menjaga konsistensi linguistik. Serta, perihal data duplikat dan data yang tingkat lengkap, data-data tersebut tidak digunakan pada fase pemrosesan data. Setelah dilakukan pembersihan dan validasi data, terdapat 6386 data yang digunakan sebagai data final pada penelitian ini. Sebelum memasuki tahapan analisa, seluruh data memasuki tahapan pemrosesan untuk memastikan data konsistensi dan kejelasan kontekstual. Mengikuti tahapan yang dijelaskan oleh Chai (2023), tahapan pemrosesan data meliputi kegiatan pemrosesan seperti yang disertakan pada Tabel 1.

Tabel 1. Tahapan Pemrosesan Data

Tahapan	Deskripsi
<i>Cleaning</i>	Penghilangan emoji, karakter khusus atau simbol (contoh, ?, !, &, dan %), dan alamat situs didalam data ulasan para pengguna.
<i>Normalization</i>	Mengkonversi seluruh kata menjadi huruf kecil dan memangkas <i>whitespace</i> pada data ulasan.
<i>Tokenization</i>	Memisahkan dan merubah kalimat menjadi token.
<i>Stopword removal dan lemmatization</i>	Menghilangkan kata-kata umum namun tidak informatif (contoh, <i>the, is, are</i>) dan mengubah seluruh kata menjadi kata dasar (contoh, <i>talking</i> menjadi <i>talk</i>).

Dengan tahapan yang telah dijelaskan, data ulasan para pengguna disimpan kedalam format CSV dengan pembagian nama kolom sebagai berikut: *review ID*, *rating*, *review text*, *review date*, *sentiment score*, *aspect*, dan *aspect sentiment*. Terkhusus untuk kolom *sentiment score* dan *aspect sentiment*, data diperoleh dengan menerapkan VADER. Untuk kolom *aspect*, data tersebut diperoleh dengan menerapkan ABSA. Dari pemrosesan yang dilakukan, jumlah data yang digunakan pada penelitian ini disertakan pada Tabel 2.

Tabel 2. Jumlah dan Persentase Komposisi Data Berdasarkan Rating

Rating Aplikasi	Jumlah Ulasan	Persentase Komposisi Data (%)
1	861	13.48
2	290	4.54
3	460	7.2
4	855	13.4
5	3920	61.38
Total	6386	100

2.3 Analisa Sentimen

2.3.1 *Valence Aware Dictionary and sEntiment Reasoner (VADER)*

Model analisa sentimen VADER diterapkan dalam mengkalkulasi polaritas sentimen setiap ulasan para pengguna, dimana data tersebut menghasilkan empat hasil, yaitu positif, negatif, netral, dan *compound score* yang merepresentasikan tingkat sentimen secara keseluruhan (Hutto & Gilbert, 2014). Kategori sentimen ditentukan dengan nilai *compound score* dengan tiga interval yang terstandarisasi: (1) ulasan dengan nilai ≥ 0.05 dikategorikan sebagai positif, (2) ulasan dengan nilai ≤ -0.05 dikategorikan sebagai negatif, dan (3) ulasan dengan nilai diantara -0.05 dan 0.05

dikategorikan sebagai netral. Nilai *compound score*, yang digunakan untuk memperoleh tingkat polaritas sentimen, di kalkulasi dengan menggunakan nilai sentimen ditingkatan kata (*lexicon*) yang telah disesuaikan dengan kondisi tata bahasa dan kontekstual (contoh, negasi dan tanda baca). Lalu nilai tersebut dinormalisasi menjadi skala kontinu dengan rentang -1 (sangat negatif) sampai dengan +1 (sangat positif), dimana hal tersebut dapat memberikan hasil evaluasi yang konsisten dan terinterpretasi untuk seluruh ulasan pengguna.

2.3.2 Aspect-Based Sentiment Analysis (ABSA)

Pendekatan ABSA diterapkan untuk menghubungkan setiap sentimen pengguna terhadap fitur (atau aspek) spesifik terhadap aplikasi ELSA. Mengikuti pendekatan yang diterapkan oleh Yang et al. (2025), ekstraksi aspek tersebut dilakukan dengan menggunakan dependensi spaCy dan mekanisme berbasis kondisi untuk mengidentifikasi pola dalam menemukan kata benda (*noun*) atau frasa yang merujuk kepada suatu fitur (misal, *subscription*, *feedback*, dan *learning*). Proses ini mengkuantifikasi nilai polaritas atas seluruh aspek yang teridentifikasi terhadap nilai sentimen yang diperoleh dengan menggunakan VADER. Dengan mengintegrasikan nilai sentimen dari VADER dengan aspek yang diperoleh dari ABSA, analisa yang dilakukan telah mengidentifikasi area pengembangan prioritas yang dapat mempengaruhi kepuasan dan retensi pengguna—terkhususnya terhadap aspek yang memiliki tingkat kepentingan yang tinggi namun memiliki nilai kepuasan yang rendah, dimana hal tersebut merupakan area pengembangan strategis yang dapat dipertimbangkan untuk menghasilkan peningkatan dan inovasi.

2.4 Uji Validitas dan Reliabilitas Model Analisa Sentimen

Didalam memvalidasi model analisa sentimen, van Atteveldt et al. (2021) menekankan bahwa evaluasi disarankan untuk tidak bergantung hanya kepada akurasi dari metrik algoritmik, namun tetap menggunakan komparasi sistematis dengan anotasi manual untuk menilai keselarasan antara hasil komputasi dan interpretasi manusia. Pendekatan tersebut menghasilkan validitas konstruk, memastikan bahwa emosi dan sentimen yang ditemukan merefleksikan persepsi dari para pengguna, bukan bias dari algoritmik. Dengan menerapkan prinsip tersebut kepada model VADER, hal ini memastikan bahwa nilai polaritas yang diperoleh menggambarkan ekspresi emosional dan sentimen yang terdapat pada ulasan dari para pengguna aplikasi.

Melengkapi pendekatan tersebut, Featherstone & Barnett (2020) menjelaskan bahwa validitas eksternal dapat diuji dengan cara melakukan analisa korelasi antara hasil sentimen dengan indikator independen (seperti rating aplikasi). Dimana, korelasi positif yang kuat mengindikasikan koherensi antara perhitungan polaritas sentimen dengan evaluasi pengguna. Pada penelitian ini, kedua pendekatan pengujian validitas tersebut juga diterapkan pada model ABSA, dengan fokus kepada polaritas sentimen dan konsistensi klasifikasi di setiap aspek dengan rating aplikasi.

2.5 Kerangka Outcome-Driven Innovation (ODI)

Kerangka ODI diterapkan untuk menginterpretasi sentimen pengguna terhadap aplikasi ELSA dan mentranslasi hasil tersebut menjadi target pengembangan. Berdasarkan Nam et al. (2021) dan Ribeiro et al. (2019), tiga pendekatan diterapkan pada penelitian ini: (1) *job identification*, setiap aspek yang telah di ekstrak dengan menggunakan ABSA dikategorisasi menjadi tiga dimensi (fungsional, emosional, dan sosial); (2) pemetaan hasil, proses konversi nilai nilai sentimen positif dan negatif yang diperoleh dari VADER menjadi nilai *satisfaction* dan *importance*; dan (3) perhitungan nilai *opportunity score* dengan menggunakan Persamaan (1). Semakin besar nilai

opportunity score mengindikasikan bahwa aspek tersebut merupakan kebutuhan pengguna yang belum terpenuhi, atau merupakan fitur yang dinilai penting oleh pengguna. Didalam hal ini, nilai *opportunity score* yang besar menunjukkan bahwa pengembangan atau inovasi disarankan untuk difokuskan (atau diprioritaskan) terhadap aspek tersebut untuk memaksimalkan kepuasan pengguna dan performa aplikasi.

$$opportunity = importance + importance - satisfaction \quad (1)$$

Skala yang digunakan untuk *satisfactions* dan *importance values* berada di rentang 0 sampai dengan 10 (Nam et al., 2021). Secara spesifik, dari sisi *importance values*, semakin besar nilai tersebut maka hal tersebut menandakan bahwa aspek tersebut sering dibicarakan (atau disebutkan). Sementara, untuk *satisfactions values*, semakin besar nilai tersebut maka hal tersebut mengindikasikan bahwa para pengguna aplikasi ELSA mendapatkan hal positif (atau puas) terhadap fitur yang digunakan. Nilai *satisfactions score* dan *importance score* diperoleh dengan mengkonversi nilai yang diperoleh dari penerapan VADER (i.e., *compound score*).

Dalam rangka memperoleh nilai *satisfaction value* dan *importance value* yang digunakan pada kerangka ODI, nilai *compound score* (berada dalam rentang -1 sampai dengan 1) yang diperoleh dari setiap ulasan pengguna dikonversi kedalam bentuk yang sesuai untuk skala 0 sampai dengan 10. Untuk *satisfaction value*, *compound score* dikonversi dengan menggunakan Persamaan (2). Didalam hal ini, nilai -1 pada *compound score* menjadi 0 (yang menandakan bahwa para pengguna sangat tidak puas terhadap suatu fitur); sementara, nilai 1 menjadi 10 (yang menandakan bahwa pengguna sangat puas terhadap suatu fitur).

$$satisfaction = \frac{compound\ score + 1}{2} \times 10 \quad (2)$$

Untuk *importance value*, nilai yang digunakan mengacu kepada frekuensi aspek yang dirujuk oleh para pengguna aplikasi ELSA. Jumlah frekuensi terhadap suatu aspek merefleksikan seberapa penting suatu fitur antar satu pengguna dengan pengguna lainnya. Persamaan (3) digunakan untuk mengkonversi *compound score* menjadi *importance value*. Nilai *importance score* dihitung dengan membagi jumlah suatu fitur disebutkan dengan total frekuensi atas aspek yang telah teridentifikasi. Hasil tersebut kemudian dipetakan dengan skala 0 sampai dengan 10 dengan menerapkan pendekatan *min-max normalization*. Dimana, f_i merujuk kepada frekuensi suatu fitur, f_{min} merujuk kepada jumlah frekuensi paling sedikit disebutkan, dan f_{max} mengacu kepada aspek dengan frekuensi paling banyak.

$$importance = 10 \times \frac{f_i - f_{min}}{f_{max} - f_{min}} \quad (3)$$

3. PEMBAHASAN

3.1 Hasil Analisa VADER dan ABSA

Dari penerapan VADER terhadap klasifikasi sentimen atas fitur aplikasi ELSA dari para pengguna, disertakan pada Tabel 3, ditemukan bahwa 86.23% (dari 6386 data) ulasan yang diberikan diberikan oleh pengguna berada dikategori positif, 9.68% negatif, dan 4.09% netral.

Tabel 3. Hasil Analisa Sentimen VADER

Sentimen Ulasan VADER	Jumlah Data	Persentase Komposisi
<i>Positive</i>	5507	86.23
<i>Negative</i>	618	9.68
<i>Neutral</i>	261	4.09

Distribusi sentimen yang ditunjukkan oleh mayoritas pengguna ELSA mengekspresikan opini positif dan kepuasan terhadap aplikasi, terkhususnya terhadap fitur yang mengacu kepada pelatihan pelafalan dan latihan bahasa yang memenuhi ekspektasi mereka. Namun, terdapat sebagian kecil pengguna yang menunjukkan sentimen negatif, mengindikasikan adanya keterbatasan pada beberapa area seperti umpan balik pembelajaran dan evaluasi performa. Berdasarkan analisa VADER (Tabel 4), rata-rata rating aplikasi ialah 4.05 (standard deviasi = 1.44) dengan nilai *compound score* sebesar 0.55 (standard deviasi = 0.41), yang berada pada interval -0.96–0.99, nilai ini merefleksikan polaritas positif (> 0.4) secara keseluruhan. Hasil ini, konsisten dengan temuan yang terdapat pada Tabel 3, mengkonfirmasi bahwa ELSA secara umum diterima oleh pengguna. Serta, dengan adanya keselarasan yang kuat antara polaritas sentimen dengan rating aplikasi, hal tersebut mendukung reliabilitas dari model VADER, dimana klasifikasi sentimen yang dilakukan merefleksikan evaluasi pengguna atas fiturnya.

Tabel 4. Statistik Deskriptif Analisa Sentimen VADER

Variable	count	mean	std	min	25%	50%	75%	max
Rating	6386	4.05	1.44	1.00	3.00	5.00	5.00	5.00
Compound Score	6386	0.55	0.41	-0.96	0.44	0.69	0.84	0.99

Dari sisi model ABSA, terdapat 5258 entitas aspek yang terdapat didalam ulasan para pengguna aplikasi. Dimana, setiap aspek tersebut merepresentasikan komponen (atau fitur) spesifik yang disebutkan oleh para pengguna (contoh, pelafalan kata, langganan, dan stabilitas sistem). Dari hasil analisa sentimen dengan model ABSA—

Tabel 5—ditemukan bahwa 79.67% (dari 5258 data) memiliki sentimen positif, sementara sentimen negatif berada di 15.84% dan netral di 4.49%. Temuan ini mengindikasikan bahwa para pengguna aplikasi ELSA, secara mayoritas, merasa terpenuhi dan puas terhadap fitur aplikasi. Namun, dengan adanya 15.84% sentimen negatif didalam ulasan tersebut, hal ini menunjukkan bahwa terdapat beberapa fitur (seperti, stabilitas aplikasi dan fitur langganan) yang membutuhkan pengembangan lebih lanjut karena belum sesuai atau memenuhi ekspektasi para pengguna aplikasi.

Tabel 5. Hasil Analisa Sentimen ABSA

Sentimen Ulasan ABSA	Jumlah Data	Persentase Komposisi
Positive	4189	79.67
Negative	833	15.84
Neutral	236	4.49

Berdasarkan nilai dari statistik deskriptif dari model ABSA, yang disertakan pada Tabel 6, ditemukan bahwa nilai rata-rata dari rating aplikasi berada di nilai 3.61 dengan standard deviasi sebesar 1.62. Nilai tersebut mengindikasikan bahwa terdapat evaluasi positif dari sisi aspek, namun nilai tersebut sedikit lebih rendah dari nilai rata-rata yang didapat dari penerapan model VADER. Secara nilai *aspect compound score*, nilai rata-rata berada di 0.49 dengan standard deviasi 0.49, dan nilai tersebut berada di rentang -0.96 sampai dengan 0.99. Temuan yang diperoleh dari statistik deskriptif tersebut menunjukkan bahwa, dari sisi analisa aspek aplikasi secara keseluruhan, status sentimen berada di posisi positif. Walaupun hasil yang diperoleh dinilai sama dengan hasil yang didapat dengan menggunakan VADER, model ABSA dapat menemukan polaritas sentimen yang terdapat didalam fitur aplikasi ELSA. Beberapa aspek tersebut ditemukan memiliki sentimen positif yang sangat tinggi (contoh, pelafalan), sementara beberapa

aspek ditemukan memiliki evaluasi negatif yang mempengaruhi pengguna dalam mempelajari bahasa asing (contoh, performa atau stabilitas aplikasi).

Tabel 6. Statistik Deskriptif Analisa Sentimen ABSA

Variable	count	mean	std	min	25%	50%	75%	max
Rating	5258	3.61	1.62	1.00	2.00	4.00	5.00	5.00
Aspect Compound	5258	0.49	0.49	-0.96	0.25	0.68	0.85	0.99

Secara pengujian validitas dan reliabilitas model VADER dan ABSA, analisa dengan menggunakan korelasi Pearson pada Tabel 7 menunjukkan bahwa terdapat korelasi yang kuat antara model ABSA dan VADER secara nilai *compound score*—nilai korelasi (*r*) berada di posisi 0.98; hal ini menunjukkan bahwa kedua model tersebut memiliki validitas ditingkat konstruk. Nilai korelasi tersebut mengindikasikan bahwa kedua model tersebut melakukan pengukuran sentimen secara konsisten walaupun terdapat perbedaan di tahapan analisa pada ulasan yang digunakan. Secara spesifik, ABSA melakukan pengukuran sentimen ditingkat aspek aplikasi, sementara VADER melakukan pengukuran ditingkat *text* (atau ulasan) secara keseluruhan.

Tabel 7. Hasil Pengujian Validitas Model VADER dan ABSA

Variable	Pearson Correlation Coefficient (<i>r</i>)		
	Compound Score VADER	Compound Score ABSA	Rating
Compound Score VADER	1		
Compound Score ABSA	0.98	1	
Rating	0.56	0.55	1

Dari nilai korelasi antara rating aplikasi dan *compound score* dari model ABSA dan VADER, ditemukan bahwa nilai korelasi antara rating dan *compound score* dari model VADER berada di nilai 0.56. Sementara, nilai korelasi antara rating aplikasi dengan *compound score* dari model ABSA berada di nilai 0.55. Nilai korelasi dari yang diperoleh dari dua model tersebut dengan rating aplikasi menunjukan tingkat korelasi yang moderat. Hal ini mengindikasikan bahwa kenaikan nilai rating akan meningkatkan nilai sentimen ke arah positif. Hubungan korelasi tersebut mengkonfirmasi validitas dari sisi eksternal untuk kedua model. Dengan kata lain, hasil sentimen yang diperoleh konsisten dengan sentimen pengguna yang direfleksikan dengan nilai rating aplikasi.

3.2 Pemetaan Prioritas Pengembangan Aplikasi ELSA Berdasarkan ODI

Analisa dengan kerangka ODI diterapkan dalam rangka mengidentifikasi aspek (atau fitur) dalam aplikasi ELSA yang mempunyai potensi paling besar untuk dikembangkan lebih lanjut berdasarkan ulasan dan sentimen dari para pengguna aplikasi—analisa menggunakan data yang sebelumnya telah diperoleh dengan menggunakan model ABSA. Tabel 8 merangkum hasil statistik deskriptif atas tiga variabel yang digunakan pada kerangka ODI, yaitu nilai *satisfactions*, *importance*, dan *opportunities*, atas 12 aspek yang telah teridentifikasi pada aplikasi ELSA. Nilai rata-rata kepuasan (*satisfactions*) berada di nilai 7.28 (dengan standard deviasi sebesar 0.28) mengindikasikan bahwa para pengguna secara mayoritas fitur yang terdapat pada aplikasi ELSA dinilai positif. Sementara, dari sisi kepentingan (*importance*), nilai tersebut berada di angka 2.68 dengan standard deviasi sebesar 2.72, hal ini mengindikasikan bahwa terdapat variabilitas terhadap seberapa sering suatu fitur disebutkan oleh para pengguna. Di sisi lain, hal ini menunjukkan bahwa terdapat beberapa fitur yang dinilai merupakan hal utama didalam pengalaman pengguna pada saat menggunakan aplikasi. Untuk nilai rata-rata *opportunities*, nilai tersebut berada di -1.92 dengan standard deviasi 5.26. Hal ini mengindikasikan bahwa, secara keseluruhan, aplikasi ELSA memiliki fitur yang

bekerja sesuai dengan ekspektasi dan memenuhi kebutuhan para pengguna, walaupun terdapat beberapa aspek yang mempunyai potensi untuk dikembangkan lebih lanjut.

Tabel 8. Statistik Deskriptif *Satisfactions*, *Importance*, dan *Opportunities* ODI Aplikasi ELSA

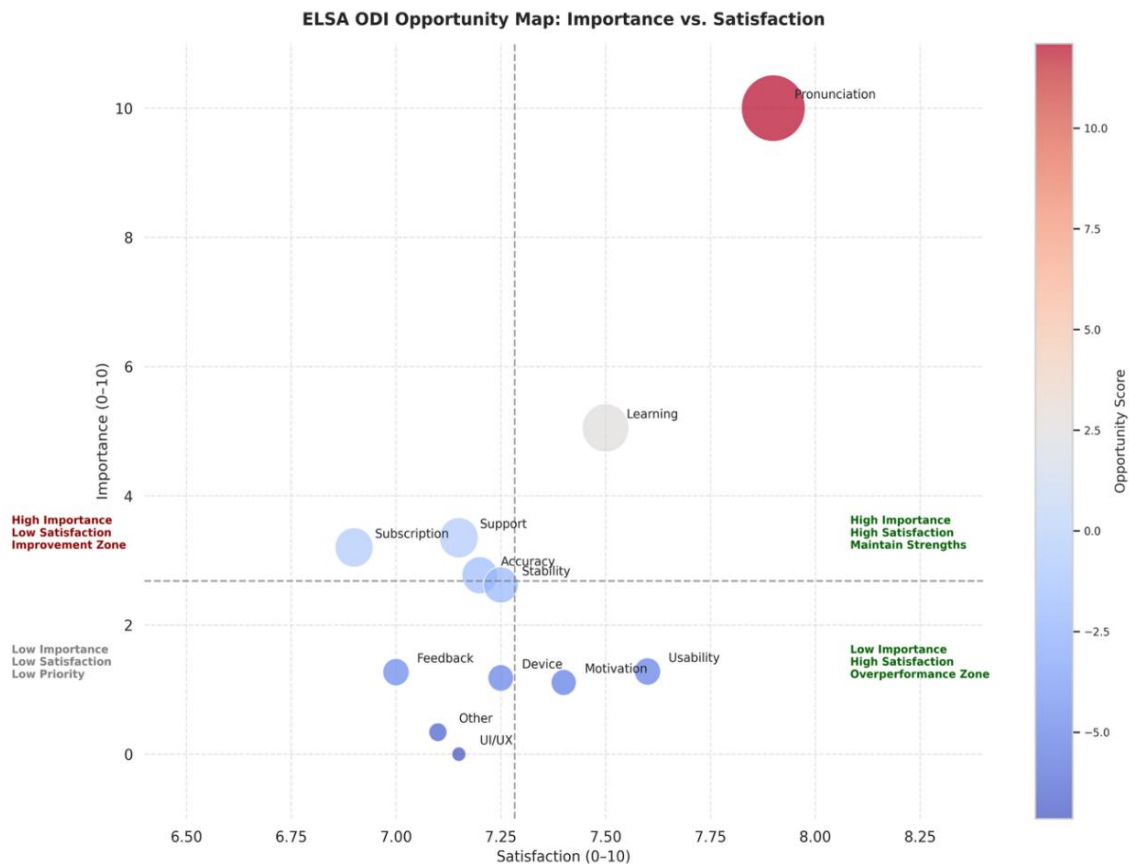
Variable	Count	Mean	Standard Deviation	Min	25%	50%	75%	Max
Aspect Frequency	12	438	373.93	69	229.5	337	515	1445
Average Compound Score	12	0.46	0.06	0.38	0.43	0.44	0.48	0.58
Satisfaction	12	7.28	0.28	6.9	7.14	7.22	7.43	7.9
Importance	12	2.68	2.72	0	1.16	1.95	3.24	10
Opportunity	12	-1.92	5.26	-7.15	-5.08	-3.24	-0.48	12.1

Tabel 9 menunjukkan hasil kalkulasi nilai *satisfaction*, *importance*, dan *opportunity* untuk setiap aspek yang terdapat didalam aplikasi ELSA, sementara Gambar 2 merupakan visualisasi atas distribusi atas ketiga nilai tersebut. Hasil menunjukkan kelebihan dan area yang dapat dikembangkan dari aplikasi tersebut. Aspek pronunciation memiliki nilai *opportunity* tertinggi (12.1) dengan nilai *importance* maksimum dan *satisfaction* paling besar (7.9), mengindikasikan area prioritas pengembangan utama terutama dalam peningkatan akurasi pelafalan dan efektivitas pembelajaran. Aspek *learning* menduduki posisi kedua dengan nilai *opportunity* moderat (2.6), menunjukkan potensi pengembangan yang mengacu kepada pembelajaran interaktif dan terpersonalisasi. Sebaliknya, aspek seperti *support*, *subscription*, dan *stability* mempunyai nilai *opportunity* yang rendah, hal ini menunjukkan bahwa para pengguna menilai bahwa aspek tersebut sudah seimbang antara nilai *satisfaction* dan *importance*. Sementara untuk aspek yang memiliki nilai *opportunity* dibawah -2 (contoh, *feedback* dan UI/UX) dinilai sebagai aspek non-kritikal. Berdasarkan Gambar 2 secara nilai *opportunity*, dapat disimpulkan bahwa aspek *pronunciation* dan *learning* merupakan aspek yang membutuhkan perhatian khusus dan pengembangan lebih lanjut (i.e., *high priority / improvement zone*). *Usability* dan *motivation* merupakan aspek yang telah memenuhi ekspektasi pengguna dan disarankan untuk dipertahankan. Sementara, untuk aspek seperti *feedback* dan UI/UX merupakan aspek yang dapat dikembangkan secara inkremental.

Tabel 9. Hasil Kalkulasi Nilai *Satisfactions*, *Importance*, dan *Opportunities* Aplikasi ELSA

Aspect	Frequency	Average Compound Score	Satisfaction	Importance	Opportunity
Pronunciation	1445	0.58	7.9	10	12.1
Learning	764	0.5	7.5	5.05	2.6
Support	530	0.43	7.15	3.35	-0.45
Subscription	510	0.38	6.9	3.2	-0.49
Accuracy	450	0.44	7.2	2.77	-1.66
Stability	429	0.45	7.25	2.62	-2.02
Feedback	244	0.4	7	1.27	-4.46
Device	232	0.45	7.25	1.18	-4.88
Usability	245	0.52	7.6	1.28	-5.04
Motivation	222	0.48	7.4	1.11	-5.18
Other	116	0.42	7.1	0.34	-6.42

Aspect	Frequency	Average Compound Score	Satisfaction	Importance	Opportunity
UI/UX	69	0.43	7.15	0	-7.15



Gambar 2. Pemetaan Pengembangan Fitur ELSA Berdasarkan Kerangka ODI

3.3 Pembahasan Hasil dan Peringkat Prioritas Pengembangan Aplikasi

Integrasi antara ABSA, VADER, dan kerangka ODI memberikan pemahaman yang komprehensif perihal sentimen dan persepsi pengguna terhadap aplikasi ELSA. Dengan mengkombinasikan polaritas sentimen, kepentingan suatu fitur, dan nilai *opportunity*, pendekatan tersebut memberikan evaluasi multidimensi atas pengalaman pengguna, dan mengidentifikasi area pengembangan prioritas dalam meningkatkan utilitas dan performa aplikasi. Secara general, temuan mengindikasikan bahwa aplikasi ELSA secara efektif memenuhi ekspektasi pada mayoritas aspek yang telah teridentifikasi. Namun, terdapat beberapa fitur yang memiliki potensi untuk optimisasi dan inovasi.

Pronunciation ditemukan sebagai aspek yang merupakan prioritas pengembangan utama dengan nilai *opportunity* tertinggi (12.1), nilai *importance* tertinggi (10), dan nilai kepuasan yang signifikan (7.9). Temuan ini mengindikasikan bahwa aspek *pronunciation* ialah nilai proposisi utama pada aplikasi ELSA dan diharapkan untuk dikembangkan secara berkelanjutan untuk membantu para pengguna dalam meningkatkan kemampuan berbahasa mereka. Temuan ini selaras dengan Liu (2023), yang menekankan bahwa keahlian dalam pelafalan merupakan hal fundamental dalam membangun kepercayaan diri pengguna dalam berbahasa asing. Sementara, aspek *learning* menduduki posisi kedua (*opportunity* = 2.6; *satisfaction* = 7.5; *importance* = 5.05), merefleksikan adanya kebutuhan pengguna atas konten pembelajaran yang adaptif dan terpersonalisasi. Penelitian yang dilakukan oleh Kim (2021), Corral Abad et al. (2021), dan Valtonen et al. (2021) menunjukkan bahwa pembelajaran adaptif dan penerapan mekanisme

pembelajaran secara *feedback-driven* meningkatkan pemahaman dan *engagement* para pengguna, dimana hal tersebut merupakan salah satu fitur yang berada pada area pengembangan strategis.

Untuk aspek lainnya—seperti *support*, *subscription*, *accuracy*, dan *stability*—menunjukkan hasil yang seimbang antara nilai *satisfaction* dan *importance*. Hal ini menghasilkan nilai negatif dari sisi *opportunity score*, mengindikasikan bahwa aspek tersebut secara umum telah memenuhi ekspektasi pengguna. Aspek tersebut tidak berada dalam tahapan untuk dilakukannya pengembangan mayor, namun perlu di supervisi secara periodik untuk menjaga tingkat reliabilita aplikasi secara berkelanjutan. Disisi lain, untuk aspek seperti *usability*, *motivation*, dan *feedback*, menunjukkan nilai *importance* yang rendah; hal ini dapat diartikan bahwa aspek tersebut bukan faktor utama dalam kepuasan pengguna. Namun, pengawasan secara periodik disarankan untuk tetap dilakukan sehubungan adanya kompetisi antar aplikasi (Martens & Maalej, 2019).

4. KESIMPULAN

Penelitian ini mengintegrasikan kerangka ODI dengan VADER dan ABSA untuk mengevaluasi ulasan pengguna dari aplikasi ELSA, dan menentukan prioritas pengembangan fitur berdasarkan nilai *satisfaction*, *importance*, dan *opportunity*. Dengan menerapkan *mixed-method design* yang mengkombinasikan analisa komputasi sentimen dengan evaluasi interpretatif, ulasan pengguna dapat dikuantifikasi dan ditransformasi menjadi tahapan pengembangan. Pendekatan tersebut bertujuan untuk mengidentifikasi sentimen paling dominan pada aplikasi ELSA, area prioritas utama yang membutuhkan pengembangan, beserta fitur yang memiliki dampak paling signifikan terhadap kepuasan pengguna apabila dikembangkan.

Hasil yang diperoleh VADER mengindikasikan bahwa 86.23% pengguna dari aplikasi ELSA mengekspresikan sentimen positif, hal ini merefleksikan kepuasan secara menyeluruh dengan fitur pembelajarannya. Temuan tersebut didukung dari temuan yang diperoleh dari ABSA, terdapat 12 aspek utama yang mempengaruhi pengalaman pengguna dalam menggunakan aplikasi. Dimana, aspek *pronunciation* dan *learning* mempunyai tingkat prioritas paling signifikan. Terdapat korelasi yang kuat ($r = 0.98$) antara VADER dan ABSA yang mengkonfirmasi validitas dan reliabilitas model; sementara terdapat korelasi moderat dengan rating aplikasi ($r \approx 0.55-0.56$) yang mengkonfirmasi validitas eksternal. Dengan menggunakan kerangka ODI, *pronunciation* ditemukan memiliki prioritas pengembangan tertinggi (*opportunity* = 12.1), hal ini mengindikasikan bahwa terdapat ekspektasi yang tinggi terhadap pelatihan pelafalan bahasa asing. Untuk urutan kedua ditempati oleh aspek *learning* (*opportunity* = 2.6), dimana hal ini mengindikasikan terdapat kesempatan untuk pengembangan fitur pembelajaran adaptif dan interaktif. Untuk aspek lainnya, seperti *support*, *subscription*, *accuracy*, dan *stability*, menunjukkan nilai yang seimbang antara *satisfaction* dan *importance*. Sementara untuk aspek *feedback*, *UI/UX*, dan *motivation* yang memiliki nilai *opportunity* lebih rendah, mengindikasikan bahwa tiga aspek tersebut hanya membutuhkan pengembangan secara inkremental.

Secara keseluruhan, penelitian ini mengkonfirmasi bahwa fitur yang terdapat di ELSA selaras dengan ekspektasi pengguna, dan terdapat aspek, apabila dikembangkan, dapat meningkatkan kepuasan, retensi, dan loyalitas pengguna. Temuan yang telah diperoleh telah mendemonstrasikan nilai praktis dari integrasi kerangka ODI dengan analisa sentimen yang dapat mengubah ulasan kualitatif pengguna menjadi prioritas pengembangan yang telah terkuantifikasi. Untuk penelitian selanjutnya, penerapan NLP tingkat lanjut (misal, *transformer-based models* seperti BERT) dapat diterapkan untuk

meningkatkan akurasi klasifikasi ABSA. Terutama dalam hal mendeteksi aspek dan pemahaman sentimen data bahasa yang telah terkontesktualisasi—dibandingkan klasifikasi berbasis *lexicon* seperti VADER.

REFERENCES

- Awajan, I., Mohamad, M., & Al-Quran, A. (2021). Sentiment Analysis Technique and Neutrosophic Set Theory for Mining and Ranking Big Data from Online Reviews. *IEEE Access*, 9, 47338–47353. <https://doi.org/10.1109/ACCESS.2021.3067844>
- Chai, C. P. (2023). Comparison of text preprocessing methods. *Natural Language Engineering*, 29(3), 509–553. <https://doi.org/10.1017/S1351324922000213>
- Corral Abad, E., Gómez García, M. J., Díez-Jimenez, E., Moreno-Marcos, P. M., & Castejón Sisamon, C. (2021). Improving the learning of engineering students with interactive teaching applications. *Computer Applications in Engineering Education*, 29(6), 1665–1674. <https://doi.org/10.1002/cae.22415>
- Featherstone, J. D., & Barnett, G. A. (2020). Validating Sentiment Analysis on Opinion Mining Using Self-Reported Attitude Scores. *2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS)*, 1–4. <https://doi.org/10.1109/SNAMS52053.2020.9336540>
- Fu, Y., Wang, Y., Ye, X., Wu, W., & Wu, J. (2023). Satisfaction with and Continuous Usage Intention towards Mobile Health Services: Translating Users' Feedback into Measurement. *Sustainability*, 15(2). <https://doi.org/10.3390/su15021101>
- Handayani, L. N. C., Sukerti, G. N. A., Nugroho, I. M. R. A., Wicaksana, K. A. B., & Ardika, I. W. D. (2022). Development and Implementation of a Mobile-Based Accounting Terms Application as Self-Learning Kit in English for Vocational Purposes. *Proceedings of the International Conference on Applied Science and Technology on Social Science 2021*, 179–187. <https://doi.org/10.2991/as-sehr.k.220301.030>
- Haug, S., Benke, I., & Maedche, A. (2023). Aligning Crowdsworker Perspectives and Feedback Outcomes in Crowd-Feedback System Design. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW1), 1–28. <https://doi.org/10.1145/3579456>
- Hutto, C. J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8, 216–225. <https://doi.org/10.1609/icwsm.v8i1.14550>
- Kim, M. K. (2021). A design experiment on technology-based learning progress feedback in a graduate-level online course. *Human Behavior and Emerging Technologies*, 3(5), 649–667. <https://doi.org/10.1002/hbe2.308>
- Liu, W. (2023). The theory of second language development for international students. *Journal for Multicultural Education*, 17(3), 367–378. <https://doi.org/10.1108/JME-08-2022-0106>
- Martens, D., & Maalej, W. (2019). Release Early, Release Often, and Watch Your Users' Emotions: Lessons from Emotional Patterns. *IEEE Software*, 36(5), 32–37. <https://doi.org/10.1109/MS.2019.2923603>
- Nam, S., Yoon, S., Raghavan, N., & Park, H. (2021). Identifying Service Opportunities Based on Outcome-Driven Innovation Framework and Deep Learning: A Case Study of Hotel Service. *Sustainability*, 13(1). <https://doi.org/10.3390/su13010391>
- Ribeiro, Á. H. P., Monteiro, P. R. R., & Luttembarck, L. (2019). The Use of the “Job to Be Done” methodology to identify value co-creation opportunities in the context of the Service Dominant Logic. *Brazilian Business Review*, 16(1). <https://doi.org/10.15728/bbr.2019.16.1.3>
- Serna-Herrera, A., Caicedo Rendón, O. M., & Rivera Martínez, W. (2025). An Approach for the Development and Maturation of ICT Products. *Administrative Sciences*, 15(10). <https://doi.org/10.3390/admsci15100383>

- Valtonen, T., Leppänen, U., Hyypiä, M., Kokko, A., Manninen, J., Vartiainen, H., Sointu, E., & Hirsto, L. (2021). Learning environments preferred by university students: a shift toward informal and flexible learning environments. *Learning Environments Research*, 24(3), 371–388. <https://doi.org/10.1007/s10984-020-09339-6>
- van Atteveldt, W., van der Velden, M. A. C. G., & Boukes, M. (2021). The Validity of Sentiment Analysis: Comparing Manual Annotation, Crowd-Coding, Dictionary Approaches, and Machine Learning Algorithms. *Communication Methods and Measures*, 15(2), 121–140. <https://doi.org/10.1080/19312458.2020.1869198>
- Yang, S., Liao, H., & Kóczy, L. T. (2025). Preference mining and fuzzy inference for hotel selection based on aspect-based sentiment analysis from user-generated content. *Journal of the Operational Research Society*, 76(7), 1414–1431. <https://doi.org/10.1080/01605682.2024.2437128>