

# Analisis Sentimen Berbasis Aspek Ulasan Pelanggan Hotel di Bali Menggunakan Metode *Decision Tree*

Ni Putu Ambalika Dewi<sup>a1</sup>, Ngurah Agus Sanjaya ER<sup>a2</sup>, AAIN Eka Karyawati<sup>a3</sup>, Ida Bagus Made Mahendra<sup>a4</sup>, Ida Bagus Gede Dwidasmara<sup>a5</sup>, I Gede Arta Wibawa<sup>a6</sup>

<sup>a</sup>Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Udayana  
Bali, Indonesia

<sup>1</sup>ambalikaambalikadewi@gmail.com

<sup>2</sup>agus\_sanjaya@unud.ac.id

<sup>3</sup>eka.karyawati@unud.ac.id

<sup>4</sup>ibm.mahendra@unud.ac.id

<sup>5</sup>dwidasmara@unud.ac.id

<sup>6</sup>gede.arta@unud.ac.id

## Abstract

The main means of tourism is the accommodation industry. Therefore, accommodation cannot be separated from the tourism industry because they both need each other. One of the accommodations that is most closely related to tourism is hospitality accommodation. With the increasing number of hotels in Bali, the hotel certainly needs the right marketing strategy. So, it is necessary to process customer reviews automatically to determine sentiment analysis based on customer reviews based on certain aspects. In this study, the author builds a model for aspect-based sentiment analysis using the Decision Tree method. The data used in this study is hotel customer review data in Indonesian language. Evaluation is done by measuring the performance of the Decision Tree model. The Decision Tree model for aspects produces performance, accuracy, precision, recall, and F1-Score, respectively 82,5%, 80%, 90,9%, and 85,1%, the Decision Tree model for service aspect sentiment produces accuracy, precision, recall performance, and F1-Score, respectively, which are 75%, 72,7%, 80%, and 76,2%, while the Decision Tree model for the sentiment of cleanliness aspect produces performance of accuracy, precision, recall, and F1-Score, respectively, which is 81,8%, 87,5%, 77,8%, and 82,4%.

**Keywords:** *Decision Tree, Confusion Matrix, TF-IDF, Aspect-Based Sentiment Analysis, Review*

## 1. Pendahuluan

Sektor pariwisata di Bali tidak dapat diragukan kemajuan dan popularitasnya. Pulau Bali mampu memikat jutaan wisatawan mancanegara ataupun domestik setiap tahunnya. Menurut data yang didapatkan dari <https://kompas.com>, pada tahun 2015, Bali menempati posisi kedua sebagai pulau dengan destinasi tujuan wisata terbaik setelah kepulauan Galapagos Ekuador. Sarana pokok kepariwisataan adalah industri akomodasi. Oleh karena itu, akomodasi sulit dipisahkan dengan industri pariwisata karena keduanya saling berkaitan. Salah satu akomodasi yang sangat erat kaitannya dengan pariwisata adalah akomodasi perhotelan. Pembangunan hotel-hotel di Bali semakin meningkat seiring dengan meningkatnya jumlah wisatawan yang berwisata ke Bali setiap tahunnya. Reservasi hotel juga semakin berkembang dengan adanya media-media digital berbasis internet yang memudahkan wisatawan untuk memesan kamar hotel tanpa perlu datang langsung ke hotel yang diinginkan. Beragamnya hotel yang ada di Bali membuat wisatawan cenderung melihat ulasan yang telah ditinggalkan oleh wisatawan sebelumnya untuk menentukan pilihan hotel yang mereka inginkan. Selain itu, harga juga menjadi faktor lain wisatawan dalam menentukan akomodasi hotel. Semakin banyaknya hotel yang ada di Bali, pihak hotel tentu membutuhkan strategi pemasaran yang tepat. Sehingga, diperlukan pengolahan ulasan pelanggan secara otomatis untuk menentukan analisis sentimen berdasarkan ulasan pelanggan berdasarkan aspek-aspek tertentu.

Sumber data yang umum digunakan untuk analisis sentimen adalah jaringan sosial yang menyimpan dan menyimpan banyak informasi. Media sosial amatlah penting sebagai sumber data yang

digunakan untuk analisis sentimen. Beberapa penelitian selanjutnya telah berkembang, dengan fokus pada pengembangan model terbaik untuk memperluas aplikasi analisis sentimen [8]. Pada penelitian sebelumnya, beberapa peneliti telah melakukan penelitian untuk memprediksi harga saham menggunakan metode *Decision Tree* dengan pembobotan TF-IDF dan TF-RF. Hasilnya menunjukkan bahwa penelitian yang dilakukan dengan TF-RF mendapatkan nilai F-Measure yang lebih besar dibandingkan dengan TF-IDF [1]. Pada penelitian lainnya, penulis melakukan penelitian untuk melihat karakteristik dari mahasiswa Universitas Cokroaminoto Palopo. Dalam penelitiannya, penulis menggunakan metode naive bayes dan metode pohon keputusan. Penelitian ini menggunakan variabel bebas. Dari penelitian yang dilakukan penulis, hasil analisis yang ditemukan oleh penulis merupakan hasil ketepatan dalam mengklasifikasi karakteristik pendaftar Universitas Cokroaminoto Palopo menggunakan metode *Naïve Bayes* sebesar 98,18% dan menggunakan metode *Decision Tree* sebesar 97,82% [2]. Dalam penelitian lain tentang klasifikasi berbasis *machine learning* menggunakan metode *decision tree*, peneliti mengeksplorasi lebih lanjut tentang metode *decision tree*. Dari studi yang dilakukan, terlihat bahwa penggunaan dataset yang berbeda mempengaruhi hasil klasifikasi menggunakan metode *decision tree* [3].

Untuk itu pada penelitian kali ini, penulis ingin mengetahui bahwa metode yang digunakan menghasilkan nilai akurasi, *precision*, *recall* dan *f-1 score* yang baik. Berdasarkan permasalahan dan penelitian-penelitian sebelumnya yang menjadi dasar untuk penelitian ini, maka penulis bermaksud untuk melakukan penelitian terhadap performa dari metode *Decision Tree* dalam analisis sentiment berbasis aspek pada ulasan pelanggan hotel di Bali dan diharapkan penelitian ini menunjukkan hasil klasifikasi yang baik.

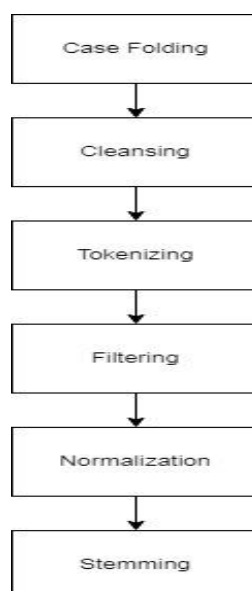
## 2. Metode Penelitian

### 2.1 Dataset

Data yang digunakan pada penelitian ini merupakan data dalam bentuk ulasan pelanggan hotel di Bali berbahasa Indonesia. Data berjumlah 800 dengan format *file \*.xlsx* yang meliputi 200 data ulasan untuk aspek pelayanan bernilai positif, 200 data ulasan untuk aspek pelayanan bernilai negatif, 200 data ulasan untuk aspek kebersihan bernilai positif dan 200 data ulasan untuk aspek kebersihan bernilai negatif.. Seluruh data sudah dilabeli ahli. Data berita kemudian dibagi menjadi dua yaitu data latih dan data uji, dengan sebanyak 80% data latih dan 20% data uji. Data latih tersebut kemudian dibagi lagi menjadi data latih dan data validasi untuk digunakan dalam proses pelatihan model dengan menggunakan *K-fold cross validation* dengan K=10.

### 2.2 Preprocessing

Preprocessing adalah proses pengolahan data yang digunakan untuk membuat format yang lebih baik. Tahapan-tahapan *preprocessing* yang dilakukan dalam penelitian ini adalah sebagai pada gambar berikut [4].

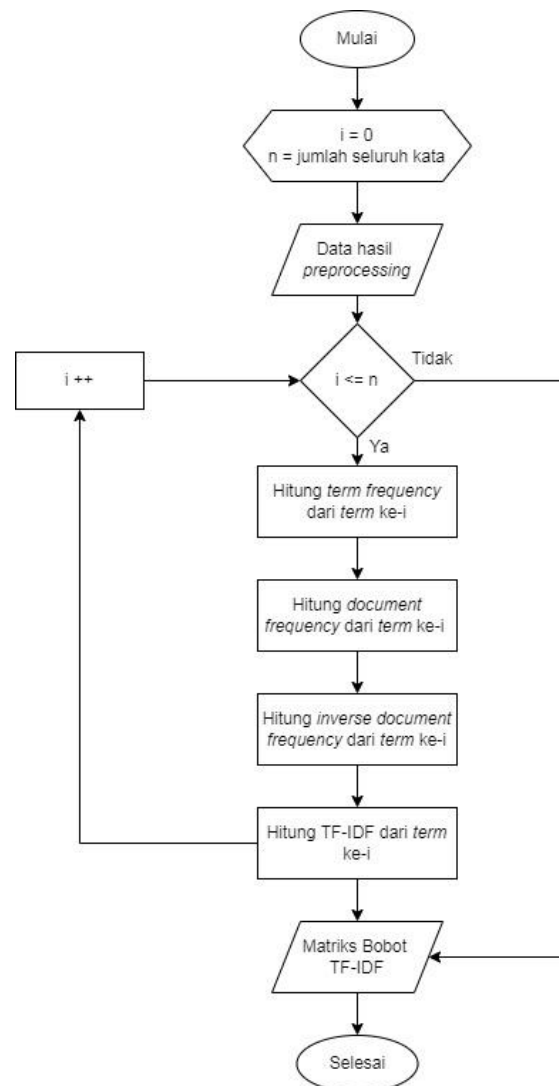


Gambar 1. Alur Preprocessing

Proses pertama pada *preprocessing* adalah *case folding* yang digunakan untuk mengubah semua huruf menjadi huruf kecil [5]. Setelah itu dilanjutkan dengan *cleansing*, pada proses ini akan dihapus karakter yang tidak memiliki kaitan terhadap analisis sentimen berbasis aspek, termasuk penghapusan tanda baca. Selanjutnya adalah *tokenizing*, yang digunakan untuk memisahkan kata dalam suatu paragraf atau kalimat menjadi token-token tertentu [5]. Selanjutnya *filtering* atau *stopwords removal*, yaitu penghapusan kata-kata yang tidak berpengaruh terhadap analisis sentimen berbasis aspek pada ulasan pelanggan [5]. Kemudian proses *normalization* yang berfungsi untuk mengubah dan mengembalikan bentuk penulisan tidak baku ke bentuk penulisan yang sesuai dengan KBBI. Proses terakhir merupakan *stemming*, yaitu proses ekstraksi kata-kata berimbuhan untuk mendapatkan kata dasar [5].

### 2.3 Term Frequency Inverse Document Frequency (TF-IDF)

Setelah melewati tahapan *preprocessing*, selanjutnya dilanjutkan dengan melakukan pembobotan dengan metode *Term Frequency Inverse Document Frequency*. Pembobotan TF-IDF dimulai dengan memasukkan data hasil *preprocessing* yang dikerjakan berulang hingga jumlah total seluruh kata. Proses selanjutnya dilakukan perhitungan *term frequency* yaitu menghitung frekuensi munculnya kata pada suatu dokumen [6]. Kemudian dilanjutkan dengan perhitungan *document frequency* untuk menghitung jumlah dokumen yang mengandung suatu kata tertentu. Selanjutnya dilanjutkan dengan perhitungan bobot *inverse document frequency* untuk menghitung distribusi kata pada koleksi dokumen dengan menggunakan persamaan (2) [6]. Terakhir, proses menghitung bobot TF-IDF dilakukan dengan mengalikan nilai *term frequency* dan nilai *inverse document frequency* dengan menggunakan persamaan (3) [6]. Gambar berikut merupakan tahapan TF-IDF.



Gambar 2. TF-IDF

a. Menghitung *term frequency* dengan persamaan

$$tf = 0,5 + 0,5 \times \frac{tf}{\max(tf)} \quad (1)$$

b. Menghitung *inverse document frequency* dengan persamaan

$$idf_t = \log\left(\frac{D}{df_t}\right) \quad (2)$$

c. Menghitung bobot TF-IDF dengan persamaan

$$W_{n,x} = tf_{n,x} \times idf_{n,x} \quad (3)$$

Keterangan:

- tf : banyaknya kata yang sering muncul pada dokumen.
- max(tf) : jumlah kata atau *term* pada data yang sama yang sering bermunculan.
- Nilai D : jumlah seluruh data yang digunakan
- df<sub>x</sub> : banyaknya data yang mengandung kata x.
- idf : banyaknya data *inverse*.
- n : data ke-n.
- x : kata ke-x.
- W<sub>d,t</sub> : bobot TF-IDF dari data ke-n dan kata ke-x.

## 2.4 Decision Tree

*Decision tree* atau bisa disebut analisis pohon keputusan adalah salah satu metode klasifikasi yang memanfaatkan teori graf dalam membagi kelompok data menjadi himpunan data. *Decision tree* secara visual mirip dengan sebuah pohon yang bercabang dengan ranting-rantingnya [7]. Data yang sudah melalui tahapan *preprocessing* dan pembobotan TF-IDF kemudian akan melalui tahap klasifikasi menggunakan metode *Decision Tree* untuk melakukan analisis sentimen berbasis aspek. Klasifikasi dilakukan dengan mengelompokkan data berupa ulasanan pelanggan yang telah di olah sebelumnya ke kelompok-kelompok aspek yang telah ditentukan. Tahap pertama yang dilakukan adalah menghitung nilai *entropy* yang akan digunakan untuk menghitung nilai *gain*. Hitung nilai *gain*, dan pilih atribut sebagai *root* berdasarkan nilai *gain* yang tertinggi. Buat cabang untuk masing-masing nilai *gain* kemudian temukan atribut terbaik dan split terbaik pada atribut. Selanjutnya lakukan pembagian data berdasarkan split, pohon akan bertumbuh dan kembali pada proses pembuatan cabang. Apabila sudah tidak terdapat cabang yang bisa dibuat maka proses telah selesai. Gambar 3 merupakan alur proses *Decision Tree*.

Tahapan-tahapan dari metode decision tree adalah sebagai berikut [1]:

a. Hitung nilai *entropy* menggunakan rumus berikut.

$$Entropy(S) = \sum_{i=1}^n - p_i \log_2 p_i \quad (4)$$

Keterangan:

- S : Sekelompok data latih
- n : Proporsi partisi pada S
- p<sub>i</sub> : Jumlah sampel pada kelas i

b. Hitung nilai *gain* menggunakan rumus berikut.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{S_i}{S} \times Entropy(S_i) \quad (5)$$

Keterangan:

- S : Sekelompok data latih
- A : Atribut
- n : Pembagian total untuk seluruh atribut A
- S<sub>i</sub> : Pembagian total ke-i

- c. Tentukan atribut sebagai *root*
- d. Proses partisi akan berhenti jika kondisi berikut terpenuhi, yaitu:
- Seluruh data yang ada pada atribut A mendapatkan kategori yang sama
  - Atribut yang ada dalam data sudah habis terbagi
  - Tidak ada data yang tersisa pada cabang yang kosong



**Gambar 3.** Decision Tree

## 2.5 Evaluasi

Evaluasi yang digunakan untuk menyelesaikan penelitian ini adalah *confusion matrix* yang dimaksudkan untuk menghitung akurasi, *precision*, *recall* dan *F1-score*. Evaluasi ditujukan untuk mengukur performa model terbaik yang sudah dipilih dari proses validasi. Ukuran evaluasi yang digunakan adalah akurasi, *precision*, *recall*, dan *F1-Score*, dan akurasi. Akurasi digunakan untuk mengevaluasi banyaknya label prediksi yang sesuai dengan label aktual, dihitung menggunakan persamaan (6). *Precision* mengukur presentase dokumen bernilai positif benar di antara seluruh dokumen yang diidentifikasi positif, dihitung menggunakan persamaan (7). *Recall* mengukur presentase dokumen bernilai positif benar yang dapat diidentifikasi di antara seluruh dokumen yang relevan, dihitung menggunakan persamaan (8). *F1-Score* adalah kombinasi hasil dari *precision* dan *recall*, dihitung menggunakan persamaan (9).

**Tabel 1. Confusion Matrix**

Data	Nilai Sesungguhnya	
	Relevan	Tidak Relevan
Retrieved	TP	FP
Not Retrieved	FN	TN

Keterangan:

TP (*True Positive*) : Proyeksi data yang benar sepenuhnya benar

FN (*False Negative*) : Terdapat data yang salah pada proyeksi data yang benar

FP (*False Positive*) : Terdapat data yang benar pada proyeksi data yang salah

TN (*True Negative*) : Proyeksi data yang salah sepenuhnya salah

Rumus berikut digunakan untuk menghitung akurasi, *precision*, *recall* dan *F-1 score* [1].

$$\text{Akurasi} = \frac{TP+TN}{(TP+FP+TN+FN)} \quad (6)$$

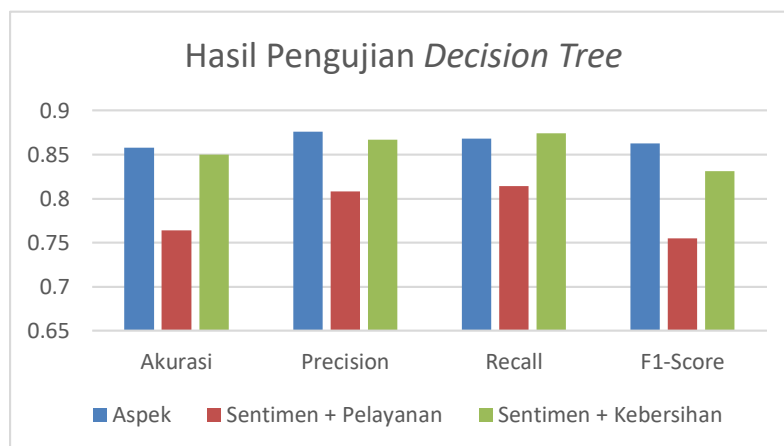
$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (7)$$

$$\text{Recall} = \frac{TP}{(TP+FN)} \quad (8)$$

$$\text{F-1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{(\text{Precision}+\text{Recall})} \quad (9)$$

## 3. Hasil dan Pembahasan

Pada pengujian penelitian ini, metode *Decision Tree* diuji untuk dapat menghasilkan nilai akurasi, *precision*, *recall* dan *F1-Score*. Proses pelatihan dan validasi menggunakan metode *K-Fold Cross Validation* untuk dapat menghasilkan performa akurasi terbaik. Pengujian dilakukan untuk membuat tiga model yaitu model aspek, model sentimen aspek pelayanan dan model sentimen aspek kebersihan. Setelah melakukan pengujian dengan menggunakan 10 *fold* untuk membuat model aspek, didapatkan nilai akurasi rata-rata sebesar 85,8%, nilai *precision* rata-rata sebesar 87,6%, nilai *recall* rata-rata sebesar 86,8%, dan nilai *F1-Score* rata-rata sebesar 86,2%. Setelah melakukan pengujian dengan menggunakan 10 *fold* untuk membuat model sentimen aspek pelayanan, didapatkan nilai akurasi rata-rata sebesar 76,4%, nilai *precision* rata-rata sebesar 80,8%, nilai *recall* rata-rata sebesar 81,4%, dan nilai *F1-Score* rata-rata sebesar 75,5%. Setelah melakukan pengujian dengan menggunakan 10 *fold* untuk membuat model sentimen aspek kebersihan, didapatkan nilai akurasi rata-rata sebesar 85%, nilai *precision* rata-rata sebesar 86,7%, nilai *recall* rata-rata sebesar 87,4%, dan nilai *F1-Score* rata-rata sebesar 83,1%. Setelah didapatkan model terbaik, selanjutnya model akan digunakan pada proses pengujian data uji dengan menggunakan data baru yang belum digunakan pada saat proses pelatihan dan validasi. Gambar berikut merupakan rata-rata hasil pengujian *Decision Tree* pada ketiga model.



**Gambar 4.** Hasil Pengujian Decision Tree

Setelah melakukan pengujian terhadap ketiga model dengan menggunakan data baru, hasil perbandingan performa model pada saat *training* dan *testing* dapat dilihat pada tabel .

**Tabel 2.** Hasil Pengujian Model

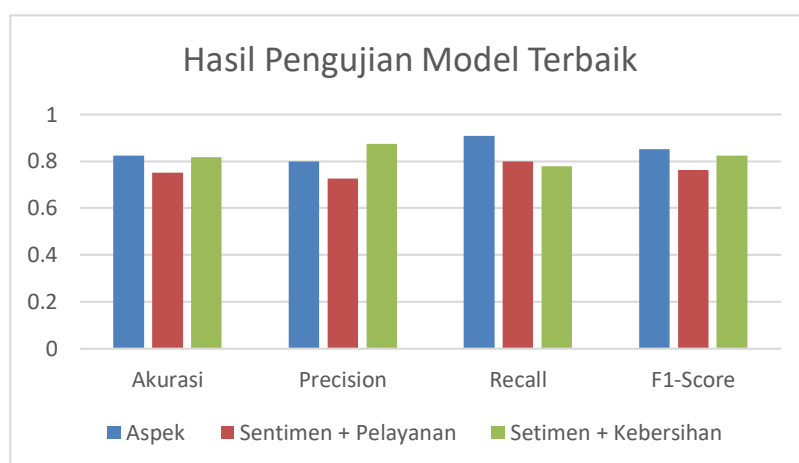
Model	Ukuran Evaluasi	Pengujian	
		Training Validasi	Testing Data Baru
Aspek	Akurasi	0.858	0.825
	Precision	0.876	0.8
	Recall	0.868	0.909
	F1-Score	0.862	0.851
Sentimen + Pelayanan	Akurasi	0.764	0.75
	Precision	0.808	0.727
	Recall	0.814	0.8
	F1-Score	0.755	0.762
Sentimen + Kebersihan	Akurasi	0.85	0.818
	Precision	0.867	0.875
	Recall	0.874	0.778
	F1-Score	0.831	0.824

Pada Gambar dapat dilihat bahwa implementasi metode *Decision Tree* dalam analisis sentimen berbasis aspek menghasilkan nilai akurasi sebesar 82,5% pada model aspek, 75% pada model sentimen aspek pelayanan dan 81,8% pada model sentimen aspek kebersihan. Akurasi yang tinggi menunjukkan bahwa model dapat dengan baik mengklasifikasikan ulasan pelanggan ke dalam aspek pelayanan dan kebersihan serta mengklasifikasikan ulasan pelanggan ke sentimen positif dan negatif. Selanjutnya, perlu adanya ukuran evaluasi lain yang memperhitungkan kesalahan prediksi, yaitu *precision*, *recall*, dan *F1-Score*.

Nilai *precision* yang dihasilkan sebesar 80% pada model aspek, 72,7% pada model sentimen aspek pelayanan dan 87,5% pada model sentimen aspek kebersihan. Nilai *precision* yang tinggi menunjukkan bahwa model dapat dengan baik mengklasifikasikan aspek pelayanan ke dalam aspek pelayanan dan sentimen positif ke dalam sentimen positif, serta sedikit kesalahan prediksi aspek pelayanan ke dalam aspek kebersihan dan sentimen positif ke dalam sentimen negatif.

Nilai *recall* yang dihasilkan sebesar 90,9% pada model aspek, 80% pada model sentimen aspek pelayanan dan 77,8% pada model sentimen aspek kebersihan. Nilai *recall* yang tinggi menunjukkan bahwa model dapat dengan baik mengklasifikasikan aspek pelayanan ke dalam aspek pelayanan dan sentimen positif ke dalam sentimen positif, serta sedikit kesalahan prediksi aspek kebersihan ke dalam aspek pelayanan dan sentimen negatif ke dalam sentimen positif.

Nilai *F1-Score* yang dihasilkan sebesar 85,1% pada model aspek, 76,2% pada model sentimen aspek pelayanan dan 82,4% pada model sentimen aspek kebersihan. Nilai *F1-Score* digunakan untuk melihat keseimbangan antara *precision* dan *recall*, sehingga nilai ini bisa digunakan juga sebagai ukuran evaluasi sebuah model selain menggunakan akurasi.



**Gambar 5.** Hasil Pengujian Model Terbaik

#### 4. Kesimpulan

Berdasarkan penelitian yang telah dilakukan, ditarik kesimpulan bahwa implementasi metode *Decision Tree* dalam analisis sentimen berbasis aspek pada ulasan pelanggan dengan menggunakan *K-Fold Cross Validation* menghasilkan tiga model terbaik. Pada pengujian data baru, model *Decision Tree* untuk aspek menghasilkan performa akurasi, *precision*, *recall*, dan *F1-Score* secara berturut-turut yaitu 82,5%, 80%, 90,9%, dan 85,1%. Pada pengujian model *Decision Tree* untuk sentimen aspek pelayanan menghasilkan performa akurasi, *precision*, *recall*, dan *F1-Score* secara berturut-turut yaitu 75%, 72,7%, 80%, dan 76,2%. Pada pengujian model *Decision Tree* untuk sentimen aspek kebersihan menghasilkan performa akurasi, *precision*, *recall*, dan *F1-Score* secara berturut-turut yaitu 81,8%, 87,5%, 77,8%, dan 82,4%.

#### Daftar Pustaka

- [1] M. G. Tambunan and E. B. Setiawan, "Prediksi Kepribadian DISC pada Twitter Menggunakan Metode *Decision Tree* C4.5 dengan Pembobotan TF-IDF dan TF-RF," *e-Proceeding of Engineering*. Vol. 7, No. 1, page 2725-2738, 2020
- [2] Y. Hastuti, "Klasifikasi Karakteristik Mahasiswa Universitas Cokroaminoto Palopo Menggunakan Metode *Naïve Bayes* dan *Decision Tree*," *Jurnal Dinamika*. Vol. 7, No.2, page 34-41, 2016
- [3] B. T. Jijo and A.M. Abdulazeez, "Classification Based on *Decision Tree* Algorithm for Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 1, page 20–28, 2021
- [4] E. Supriyanti and M. Iqbal, "Pengukuran *Similarity* Tema pada Juz 30 Al Qur'an Menggunakan Teks Klasifikasi," *Jurnal SIMETRIS*, Vol. 9, No. 1, page. 361–370, 2018



- [5] K. D. Y. Wijaya and A.A.I.N.E. Karyawati, "The Effects of Different Kernels in SVM Sentiment Analysis on Mass Social Distancing," *JELIKU*, vol. 9, no. 2, page 161-168, Nov. 2020
- [6] I. W. Santiyasa, G. P. A. Brahmantha, I. W. Supriana, I. G. G. A. Kadyanan, I. K. G. Suhartana, and I. B. M. Mahendra, "Identification of Hoax Based on Text Mining Using K-Nearest Neighbor Method," *JELIKU (Jurnal Elektron. Ilmu Komput. Udayana)*, vol. 10, no. 2, page 217–226, 2021
- [7] C. Z. Janikow, "Fuzzy decision trees: issues and methods," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*. Vol. 28, No. 1, page 1–14, 1998
- [8] N. C. Dang, M. N. Moreno-García, and F. De la Prieta, "Sentiment Analysis Based on Deep Learning: A Comparative Study," *Electronics*, vol. 9, no. 3, page 483, Mar. 2020.

*This page is intentionally left blank.*