

Analisis Sentimen Top 10 Traveler Ranked Hotel di Kota Makassar Menggunakan Algoritma Decision Tree dan Support Vector Machine

Yerik Afrianto Singgalen*

Fakultas Ilmu Administrasi Bisnis dan Ilmu Komunikasi, Pariwisata, Universitas Katolik Indonesia Atma Jaya, Jakarta, Indonesia

Email: yerik.afrianto@atmajaya.ac.id

Email Penulis Korespondensi: yerik.afrianto@atmajaya.ac.id

Abstrak—Perencanaan perjalanan wisata yang dipopulerkan oleh influencer memantik khalayak untuk meninjau destinasi berdasarkan citra dan daya tarik sebelum merancang perjalanan wisata sendiri, bersama keluarga, bersama teman, maupun bersama pasangan. Penelitian ini menawarkan gagasan untuk menganalisis sentimen wisatawan yang menggunakan layanan akomodasi hotel di kota Makassar berdasarkan data ulasan website Tripadvisor. Metode yang digunakan ialah Cross-Industry Standard Procedure for Data Mining (CRISP-DM) dengan tahapan sebagai berikut : tahap business understanding, tahap data understanding, tahap data preparation, tahap modeling, tahap evaluation, dan tahap deployment. Hasil penelitian ini menunjukkan bahwa tahap business understanding menunjukkan pentingnya identifikasi dan analisis perilaku wisatawan terkait penilaian location, cleanliness, service, dan value hotel sebagai langkah strategis untuk menyusun program pencitraan hotel agar menarik wisatawan dengan karakteristik tamu (solo, couple, business, family). Tahap data understanding merujuk pada karakteristik data yang diproses sesuai metode CRISP-DM, serta data pendukung untuk analisis dan pembahasan konteks Top 10 Hotel rekomendasi Tripadvisor di kota Makassar. Pada tahap data preparation, jumlah data teks yang telah diproses menggunakan algoritma Decision Tree (DT) dan Support Vector Machine (SVM) sebesar 1.138 melalui pre-processing data (tokenize, transform cases, filter tokens by length, stopwords, stemming). Pada tahap modeling, algoritma yang menunjukkan performa terbaik ialah SVM dengan nilai akurasi 98,98%, nilai presisi 100%, nilai recall 97,96%, nilai f-measure 98,97%, dan nilai AUC 100%. Pada tahap evaluasi, dapat diketahui bahwa klasifikasi ulasan dominan pada sentimen positif dibandingkan sentimen negatif. Dengan demikian, rekomendasi untuk tahap deployment ialah optimalisasi produk dan layanan yang berhubungan dengan room amenity, room feature, room type, cleanliness, service, value, dan location.

Kata Kunci: Hotel; Traveler Ranked; Decision Tree; Support Vector Machine; Makassar

Abstract—Travel planning popularized by influencers sparks audiences to review destinations based on image and attractiveness before designing a trip alone, with family, with friends, or as a couple. This research offers ideas to analyze the sentiment of travelers who use hotel accommodation services in Makassar based on Tripadvisor website review data. The method used is the Cross-Industry Standard Procedure for Data Mining (CRISP-DM) with the following stages: business understanding stage, data understanding stage, data preparation stage, modeling stage, evaluation stage, and deployment stage. The results of this study show that the business understanding stage shows the importance of identifying and analyzing tourist behavior related to the assessment of location, cleanliness, service, and hotel value as a strategic step to develop a hotel imaging program to attract tourists with guest characteristics (solo, couple, business, family). The data understanding stage refers to the features processed according to the CRISP-DM method and supporting data for analysis and discussion of the context of Tripadvisor's Top 10 recommended hotels in Makassar. At the data preparation stage, the amount of text data that has been processed using the Decision Tree (DT) algorithm and Support Vector Machine (SVM) is 1,138 through data pre-processing (tokenize, transform cases, filter tokens by length, stopwords, stemming). At the modeling stage, the algorithm that shows the best performance is SVM, with an accuracy value of 98.98%, a precision value of 100%, a recall value of 97.96%, an f-measure discount of 98.97%, and an AUC value of 100%. At the evaluation stage, it can be seen that the classification of reviews is dominant on positive sentiment compared to negative sentiment. Thus, the recommendation for the deployment stage is optimizing products and services related to room amenities, room features, room type, cleanliness, service, value, and location.

Keywords: Hotel; Traveler Ranked; Decision Tree; Support Vector Machine; Makassar

1. PENDAHULUAN

Sektor pariwisata di Indonesia mengalami perkembangan karena dukungan pelbagai pemangku kepentingan, salah satunya ialah industri perhotelan yang menyediakan layanan akomodasi bagi para wisatawan. Yuliamir et al. menunjukkan bahwa hotel berperan penting dalam meningkatkan kunjungan wisatawan di suatu destinasi wisata, meskipun demikian upaya untuk memantik minat, kepuasan, dan loyalitas tamu hotel sangat bergantung pada strategi masing-masing hotel [1]. Disisi lain, Runtuwarouw mengemukakan bahwa sektor pariwisata mampu menciptakan lapangan pekerjaan melalui industri perhotelan sehingga memantik kunjungan wisatawan, dan pertumbuhan sektor pendukung lainnya [2]. Hal ini berarti bahwa perkembangan industri perhotelan di Indonesia tidak terlepas dari aktivitas pariwisata yang menjadi katalisator ekonomi sekaligus mobilisasi kesejahteraan sosial masyarakat. Dengan demikian dapat diketahui bahwa kajian tentang industri perhotelan perlu dilakukan untuk mengidentifikasi dan menganalisis persepsi tamu hotel ketika menggunakan layanan akomodasi selama berwisata di berbagai destinasi wisata.

Penetapan Kawasan Strategis Pariwisata Nasional (KSPN) serta destinasi pariwisata super prioritas di Indonesia memantik partisipasi investor penyedia layanan akomodasi dan amenitas. Halum et al. menunjukkan bahwa upaya untuk mewujudkan pariwisata berkelanjutan di destinasi wisata super priortias ditempuh dengan program pendampingan masyarakat yang tinggal di sekitar destinasi wisata agar dapat memanfaatkan peluang ekonomi, serta meningkatkan kesejahteraan sosial [3]. Disisi lain, Kurniansah dan Purnama berpendapat bahwa pembangunan infrastruktur berupa hotel, resort, dan cottage meningkat secara signifikan setelah adanya penetapan dan arah kebijakan pembangunan destinasi wisata super prioritas di Indonesia [4]. Hal ini menunjukkan adanya hubungan antara pertumbuhan sektor pariwisata dan

industri perhotelan di Indonesia. Namun, kajian tentang analisis sentimen wisatawan terhadap produk dan layanan hotel di destinasi wisata prioritas maupun daerah penghubung (transit) belum dilakukan secara merata dan berkala untuk memperoleh data terbaru. Berdasarkan hal tersebut maka penelitian ini fokus menganalisis sentimen tamu hotel terhadap *Top 10 Traveler Ranked Hotel* di Kota Makassar pada platform Tripadvisor.

Kota Makassar memiliki daya tarik wisata yang beragam sekaligus berperan penting sebagai area transit yang menjadi titik sentral mobilitas wisatawan ke Indonesia bagian Timur maupun Barat. Dianindarwati et al. menunjukkan bahwa arah kebijakan dan program pengembangan pariwisata di Kota Makassar menekankan pada kolaborasi lintas sektor atau pemangku kepentingan untuk mempercepat pertumbuhan ekonomi di masing-masing obyek wisata [5]. Selain itu, Hardianti dan Nasrullah menegaskan bahwa strategi pemasaran bauran yang diterapkan menjadi faktor penarik wisatawan untuk berkunjung ke kota Makassar [6]. Selain itu, Sasmita et al. menunjukkan adanya pembaharuan prosedur check-in dan check-out di beberapa Hotel Bintang Empat Kota Makassar sebagai bentuk inovasi dan adaptasi layanan akomodasi di era *new normal* [7]. Hal ini berarti produk dan layanan akomodasi hotel berbintang di Kota Makassar diperbaharui untuk meningkatkan kenyamanan dan keamanan tamu hotel. Mempertimbangkan hal tersebut maka penelitian ini fokus pada analisis sentimen terhadap produk dan layanan Top 10 Traveler Ranked Hotel yang direkomendasikan Tripadvisor sebagai berikut : Claro Makassar; The Rinra Makassar; Swiss-Belhotel Makassar; Novotel Makassar Grand Shayla; Almadera Hotel; Dalton Hotel Makassar; MaxOneHotels @Resort Makassar; Ibis Makassar City Center; Harper Perintis-Makassar; Aryaduta Makassar.

Masing-masing hotel memiliki standar pelayanan prima yang sesuai dengan tema sebagai keunikan atau daya tarik, namun secara keseluruhan dapat diklasifikasikan berdasarkan *room amenities*, *room feature*, dan *room type*. Hattah et al. menunjukkan bahwa hotel memiliki strategi pemasaran yang dapat diklasifikasikan berdasarkan produk, harga, tempat, promosi, proses, orang, bukti fisik, dan daya tarik pelanggan [8]. Dissi lain, Arimbawa et al. berargumen bahwa komunikasi faktor yang memengaruhi motivasi dan kepuasan karyawan sehingga meningkatkan performa [9]. Hal ini berarti bahwa optimalisasi produk dan layanan hotel ditentukan oleh strategi pemasaran hotel. Adapun, Clarita menunjukkan bahwa data ulasan tamu hotel melalui laman web Tripadvisor dapat digunakan sebagai data teks untuk analisis sentimen emosional tamu [10]. Adapun, Harwani berpendapat bahwa pihak hotel menggunakan platform Tripadvisor sebagai media pemasaran sekaligus konstruksi brand atau citra hotel agar dapat menarik tamu [11]. Hal ini berarti bahwa data ulasan tamu hotel di website Tripadvisor dapat digunakan sebagai data teks untuk dianalisis secara komprehensif sehingga dapat memberikan rekomendasi bagi pihak hotel untuk mengoptimalkan manajemen produk dan layanan yang sesuai dengan preferensi tamu. Mempertimbangkan hal tersebut maka penelitian ini menggunakan data ulasan dari *Top 10 Traveler Ranked Hotel* yang direkomendasikan Tripadvisor untuk diproses menggunakan metode CRISP-DM.

Penambahan data dapat dilakukan menggunakan metode *Cross-Industry Standard Procedure for Data Mining* (CRISP-DM) melalui tahap *business understanding*, tahap *data understanding*, tahap *data preparation*, tahap *modeling*, tahap *evaluation*, tahap *deployment*. Sukmana menunjukkan bahwa CRISP-DM menjadi salah satu metode yang efektif digunakan dalam merekomendasikan program optimalisasi proses bisnis yang sifatnya kontekstual [12]. Yudiana et al. berpendapat bahwa metode CRISP-DM dapat digunakan untuk mempertahankan pelanggan [13]. Hal ini menunjukkan bahwa analisis sentimen tamu hotel terhadap produk dan layanan dapat menggunakan metode CRISP-DM untuk merekomendasikan program peningkatan kualitas produk dan layanan berdasarkan hasil klasifikasi sentimen positif dan negatif secara visual. Mempertimbangkan hal tersebut maka penelitian ini akan menggunakan instrumen berupa aplikasi Rapidminer sebagai alat untuk membersihkan dan mengolah data (*pre-processing* dan *processing*), pemodelan (*modeling*), mengevaluasi dan merekomendasikan model yang relevan dengan konteks bisnis (*evaluation* dan *deployment*). Dengan demikian, algoritma yang akan digunakan dalam proses klasifikasi sentimen ialah *Decision Tree* (DT) dan *Support Vector Machine* (SVM).

Algoritma *Decision Tree* (DT) dan *Support Vector Machine* (SVM) dapat digunakan sebagai model dalam klasifikasi sentimen. Meskipun demikian untuk merekomendasikan model yang relevan dengan konteks penelitian ini maka perlu dibandingkan nilai akurasi, presisi, *recall* dan *Area Under Curve* (AUC). Pramesti dan Pratiwi menggunakan algoritma *Decision Tree* (DT) dan *Support Vector Machine* (SVM) untuk menganalisis sentimen program MBKM [14]. Selanjutnya, Romadloni dan Supriyanti membandingkan performa algoritma *Decision Tree* (DT), *k-Nearest Neighbour*, *Naïve Bayes Classifier* (NBC) dan *Support Vector Machine* (SVM) dalam analisis sentimen penggunaan teknologi dalam pendidikan anak [15]. Hal ini berarti bahwa masing-masing algoritma yang digunakan dalam proses klasifikasi memiliki performa yang berbeda-beda, menanggapi perbedaan jumlah data teks (data latih dan data uji), label data, dan konteks pembahasan. Mempertimbangkan hal tersebut maka penelitian ini memberikan beberapa batasan algoritma yang digunakan dalam proses klasifikasi sentimen yakni algoritma *Decision Tree* (DT) dan *Support Vector Machine* (SVM) dengan pembagian data latih 30% dan data uji 70%. Dengan demikian, dapat diketahui algoritma dengan performa terbaik yang dapat dijadikan model untuk klasifikasi sentimen dengan data teks serupa.

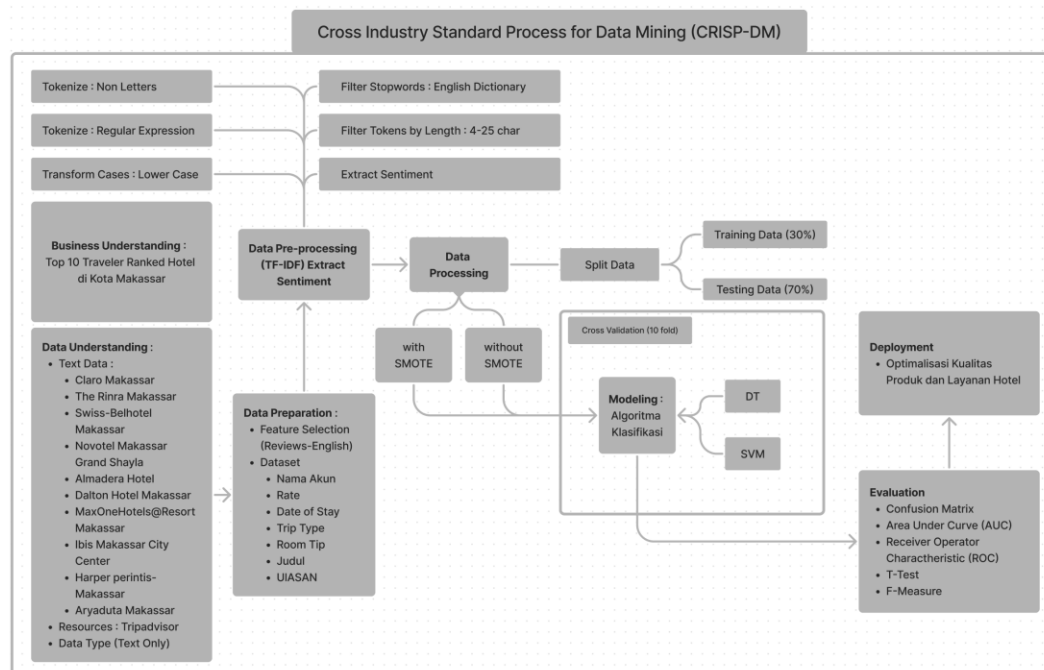
Secara teoretis, kajian ini berkontribusi pada kajian sistem informasi pariwisata yakni model penambahan dan pengolahan data digital (teks) yang berhubungan dengan perilaku wisatawan dalam penggunaan teknologi. Perilaku wisatawan ketika mengulas pengalaman menginap di hotel, menggambarkan ekspektasi terhadap produk dan layanan akomodasi hotel. Adapun, kajian tentang analisis sentimen menggunakan teknik penambahan data yang berhubungan dengan pengalaman tamu menggunakan layanan akomodasi hotel masih sangat terbatas sehingga perlu dilakukan. Adapun, secara kontekstual studi kasus *Top 10 Travel Ranked Hotel* kota Makassar dapat diproses melalui metode CRISP-DM sehingga menghasilkan data yang menggambarkan atensi tamu hotel selama menginap. Data tersebut dapat

dijadikan rekomendasi program untuk manajemen produk dan layanan agar dapat memenuhi ekspektasi atau meningkatkan kepuasan dan loyalitas tamu. Dengan demikian, luaran penelitian ini juga berkontribusi secara empiris, dimana pihak hotel dapat memanfaatkan data klasifikasi sentimen dan aspek-aspek yang menjadi atensi tamu untuk membenahi strategi pelayanan hingga pemasaran produk.

2. METODOLOGI PENELITIAN

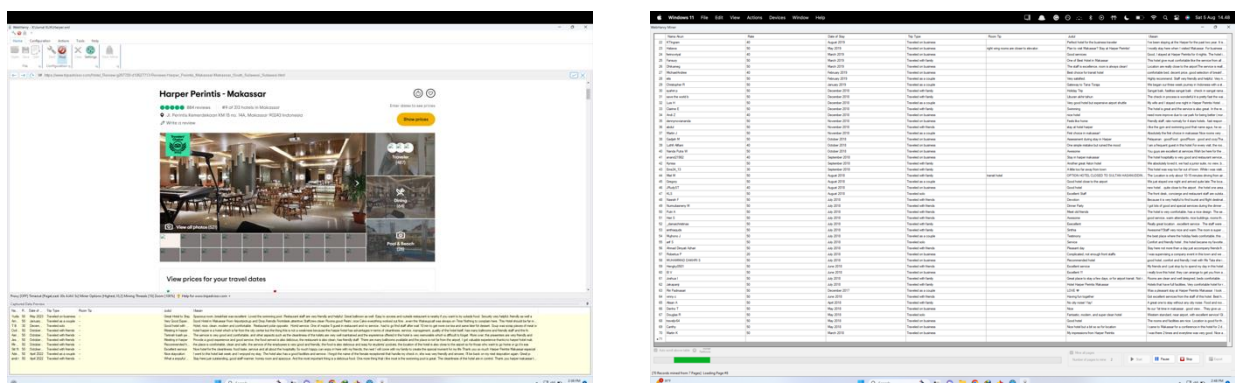
2.1 Cross-Industry Standard Procedure for Data Mining (CRISP-DM)

Metode yang digunakan dalam penelitian ini ialah metode *Cross-Industry Standard Procedure for Data Mining* (CRISP-DM) melalui tahap *business understanding*, tahap *data understanding*, tahap *data preparation*, tahap *modeling*, tahap *evaluation*, tahap *deployment*. Adapun, alur penelitian ini dapat dilihat pada gambar berikut.



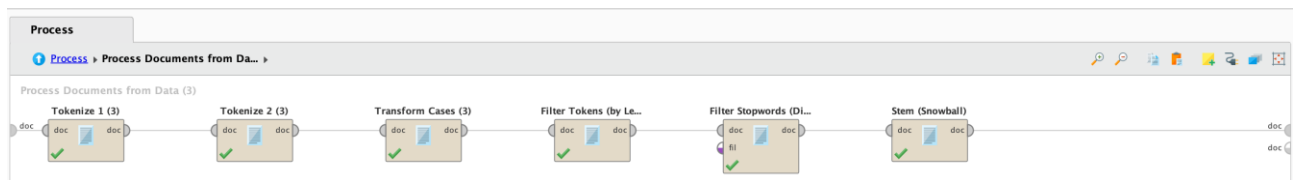
Gambar 1. Metode *Cross-Industry Standard Process for Data Mining* (CRISP-DM)

Gambar 1 merupakan proses implementasi metode CRISP-DM penelitian ini. Pada tahap *business understanding*, fokus utama dalam pembahasan penelitian ini terbatas pada *top 10 traveler ranked hotel* di Kota Makassar yang direkomendasikan oleh Tripadvisor. Adapun, pada tahap *data understanding*, data utama yang digunakan dalam analisis sentimen ialah data teks ulasan tamu hotel di kolom review dengan pengaturan *traveler rating* (*excellent, very good, average, poor, terrible*), *time of year* (*mar-may, june-aug, sep-nov, dec-feb*), *traveler type* (*families, couples, solo, business, friends*), *language* (*english*). Adapun, data-data yang dikumpulkan terbatas pada data teks untuk hotel-hotel berikut : Claro Makassar; The Rinra Makassar; Swiss-Belhotel Makassar; Novotel Makassar Grand Shayla; Almadera Hotel; Dalton Hotel Makassar; MaxOneHotels @Resort Makassar; Ibis Makassar City Center; Harper Perintis-Makassar; Aryaduta Makassar. Adapun, proses pengumpulan atau scraping data menggunakan aplikasi Webharvy sebagaimana gambar berikut ini.



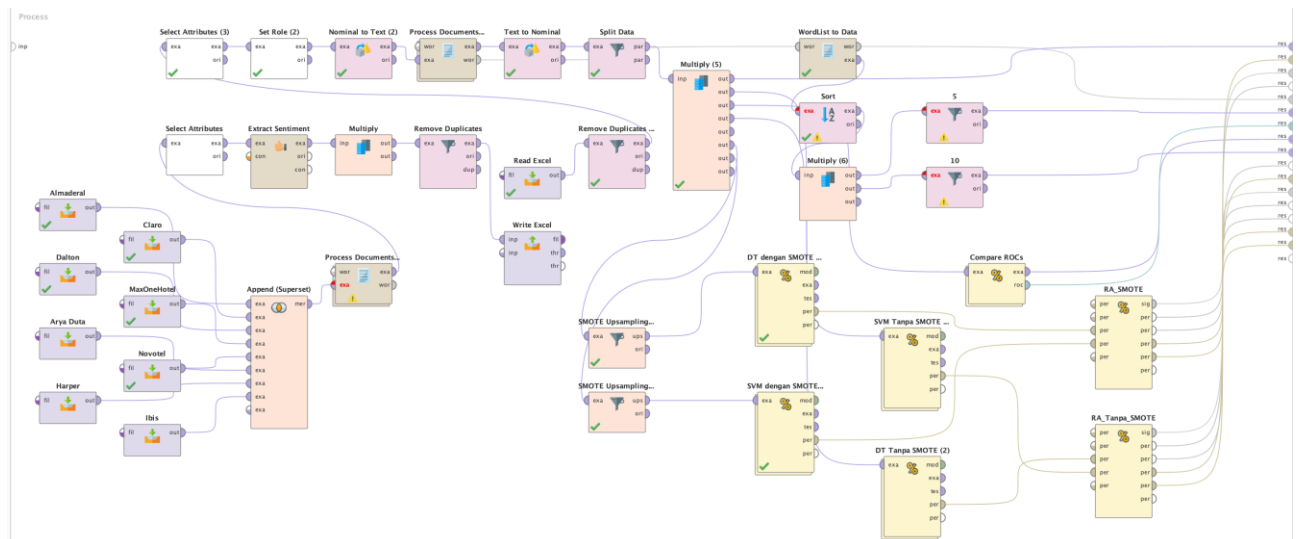
Gambar 2. Proses Pengumpulan Data (*Scraping*)

Gambar 2 merupakan proses pengumpulan data menggunakan aplikasi Webharvy. Masing-masing hotel memiliki jumlah ulasan yang berbeda, demikian halnya klasifikasi ulasan dalam bahasa yang digunakan juga berbeda-beda. Berikut ini adalah data jumlah total ulasan hotel hingga tanggal 4 Agustus 2023 : Claro Makassar (4.196); The Rinra Makassar (1.953); Swiss-Belhotel Makassar (1.769); Novotel Makassar Grand Shayla (1.682); Almadera Hotel (294); Dalton Hotel Makassar (507); MaxOneHotels @Resort Makassar (893); Ibis Makassar City Center (604); Harper Perintis-Makassar (884); Aryaduta Makassar (1.058). Meskipun demikian, karakteristik data yang digunakan dalam klasifikasi sentimen dibatasi pada data teks dalam bahasa inggris dengan pertimbangan untuk mengidentifikasi ekspektasi tamu asing atau wisatawan mancanegara ketika menggunakan layanan akomodasi dari *top 10 traveler ranked hotel* di Kota Makassar. Dengan demikian, jumlah data teks berbahasa inggris yang digunakan dari masing-masing hotel ialah sebagai berikut : Claro Makassar (336); The Rinra Makassar (261); Swiss-Belhotel Makassar (213); Novotel Makassar Grand Shayla (427); Almadera Hotel (34); Dalton Hotel Makassar (75); MaxOneHotels @Resort Makassar (67); Ibis Makassar City Center (101); Harper Perintis-Makassar (119); Aryaduta Makassar (291). Dengan demikian, data teks berbahasa inggris dapat dilanjutkan ke tahap *data preparation*. Adapun, proses pembersihan data (*data pre-processing*) dapat dilihat pada gambar berikut.



Gambar 3. Proses Pembersihan data

Gambar 3 merupakan proses pembersihan data di aplikasi Rapidminer. Pada tahap *data preparation*, jumlah data teks yang telah dikumpulkan dari *top 10 traveler ranked hotel* di Kota Makassar dari platform Tripadvisor ialah sebanyak 1.924 data ulasan berbahasa inggris. Proses labeling dilakukan dengan cara mengeliminasi *rating* netral dan menggunakan operator *extract sentiment* pada aplikasi Rapidminer untuk mendapatkan score dan label negatif dan positif. Namun, sebelum proses ekstrak data teks dilakukan, data dibersihkan terlebih dahulu menggunakan beberapa proses yakni *tokenize : non letters*, *tokenize : regular expression*, *transform cases : lower case*, *filter tokens by length : 4-25 char*, *filter stopwords : English dictionary*. Tujuan dari *data pre-processing* ialah untuk menghilangkan data duplikat, menyaring simbol dan angka dari data ulasan, mengubah huruf kapital menjadi *lower case*, serta menghilangkan kata sambung atau jenis kata yang tidak memiliki makna. Adapun, alur *data pre-processing* dan *data processing* di aplikasi Rapidminer dapat dilihat pada gambar berikut.



Gambar 4. Proses Pengolahan Data di Rapidminer

Gambar 4 merupakan alur *data pre-processing* dan *data processing* di aplikasi Rapidminer. Data teks dari masing-masing hotel dijadikan satu kemudian dibersihkan untuk diekstrak sentimen sehingga memperoleh score masing-masing kata. Selain proses klasifikasi sentimen positif dan negatif, data teks yang termasuk lima dan sepuluh kata populer akan divisualisasikan dalam bentuk *wordlist to data (bar-horizontal)*. Selanjutnya, pada tahap *modeling*, algoritma *Decision Tree* (DT) dan *Support Vector Machine* (SVM) digunakan untuk mengklasifikasi sentimen positif dan negatif. Adapun, operator *Synthetic Minority Oversampling Technique* (SMOTE) digunakan untuk mengatasi masalah keseimbangan data. Selanjutnya, pada tahap evaluasi, performa algoritma *Decision Tree* (DT) dan *Support Vector Machine* (SVM) dinilai berdasarkan perbandingan nilai akurasi, presisi, *recall*, AUC, f-measure, dan ROC. Dengan demikian, dapat direkomendasikan model yang relevan digunakan dalam proses analisis sentimen. Penelitian ini tidak hanya

memperhatikan nilai akurasi, presisi dan *recall* tetapi mempertimbangkan nilai AUC dari algoritma yang digunakan, apabila nilai AUC menunjukkan nilai performa yang lebih baik maka algoritma tersebut dapat direkomendasikan sebagai model yang tepat dalam pengolahan data teks.

2.2 Decision Tree (DT)

Decision Tree (DT) merupakan model prediksi yang menggunakan struktur pohon untuk mencari dan membuat keputusan, serta memecahkan masalah dengan mempertimbangkan berbagai faktor di dalam lingkup masalah tersebut [16]. *Decision Tree* memiliki beberapa algoritma salah satunya *Iterative Dychotomizer version* (ID3), yaitu model klasifikasi yang berupa pohon keputusan secara top-down dengan cara kerja mengevaluasi semua atribut menggunakan suatu ukuran statistik berupa *information gain* untuk mengukur efektifitas suatu atribut dalam mengklasifikasi *sample* data. Dalam algoritma ini, dibutuhkan nilai *entropy* dan *gain*, dimana *entropy* merupakan parameter untuk mengukur jumlah keberagaman atau keberadaan dalam sebuah himpunan data, sedangkan *gain* merupakan perolehan informasi sebagai ukuran efektifitas suatu atribut. Berikut adalah persamaan untuk mendapatkan nilai *entropy* dan *gain*.

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

$$Gain(S, A) = S - \sum_{i=1}^n \frac{|S_i|}{|S|} \times Si \quad (2)$$

Dimana S merupakan nilai Entropy, p_i jumlah yang memiliki nilai positif atau negatif pada kumpulan data untuk sifat tertentu. Disisi lain, Gain (S,A) adalah hasil informasi yang berasal dari luaran data yang dikelompokkan sesuai dengan atribut A. Selanjutnya, S_i adalah subset dari nilai *entropy* yang mempunyai nilai i . Adapun, S adalah subset dari nilai Entropy.

2.3 Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan salah satu algoritma yang dapat digunakan untuk klasifikasi data menggunakan hyperplane [17]. Karim menjelaskan bahwa konsep SVM menitikberatkan pada *risk minimization*, yaitu estimasi fungsi dengan cara meminimalisir batas dari *generalization error*, sehingga SVM mampu mengatasi *overfitting* [18]. Adapun, fungsi regresi dari metode SVM adalah sebagai berikut.

$$f(x) = w^T \phi(x) + b \quad (3)$$

Dimana w merupakan vector pembobot, $\phi(x)$ merupakan sebuah fungsi yang memetakan x ke dalam suatu dimensi, dan b merupakan faktor bias. Selanjutnya, Nida et al., menunjukkan bahwa SVM memiliki kelebihan dalam generalisasi data yang tinggi serta mampu menghasilkan model klasifikasi yang baik meskipun dilatih dengan data yang relatif sedikit. Meskipun demikian, sangat sulit diaplikasikan untuk himpunan data dengan sampel dan dimensi yang besar [19]. Hal ini menunjukkan bahwa SVM mampu menghasilkan performa yang baik meskipun dengan jumlah data yang relatif sedikit. Dengan demikian, perlu dilakukan pengujian algoritma SVM terhadap *top 10 traveler ranked hotel* di Kota Makassar.

3. HASIL DAN PEMBAHASAN

Penelitian ini memproses data latih (30%) dan data uji (70%) dari 1.138 data teks yang telah diberikan label negatif dan positif (binominal). Berdasarkan hasil ekstrak sentimen diperoleh 88 data teks dengan label sentimen negatif dan 1049 data teks dengan label sentimen positif. Adapun, hasil ekstrak sentimen dapat dilihat dari tabel berikut.

Tabel 1. Hasil Ekstrak Sentimen

Ulasan	Sentimen	Score	Scoring String
Okay, i stay in this hotel with my family we booked 3 rooms but the service is disappointed. First let me complain about the waitress in the restaurant. They have no manner! They interrupting people who still eating! Second, the receptionist have no coordination and not friendly. I extend the booked because I delayed my flight to be back to Jakarta and I already told a receptionist whose on morning shift but then when I back to the hotel at the afternoon, the afternoon shift receptionist said that they didn't get an information at all, and she doesn't look friendly at all! She told me to pay the deposits for my second book, when the first receptionist is already told me to pay when i check out! So when the first one said A, why the second one said B and make the customer confuse? Oh her name is Nika, the second one. Really disappointed and you have to be more friendly miss Nika! And the last one, the pool staff!!! She told me that I have to rent the swim suit if I want to swim when the other customer is wearing tshirt and short. Mister manager please train your people. Thank You for the disappointed services! Never forget this hotel!	Negatif	-1,07692307692308	okay (0.23) disappointed (-0.54) complain (-0.38) no (-0.31) interrupting (-0.31) no (-0.31) friendly (0.56) extend (0.18) delayed (-0.23) friendly (0.56) pay (-0.10) pay (-0.10) confuse (-0.23) disappointed (-0.54) friendly (0.56) miss (-0.15) want (0.08) please (0.33) thank (0.38) disappointed (-0.54) forget (-0.23)

I came to Makassar for a conference in this hotel for 2 days and 1 night. The hotel is really nice and clean. It's also really modern so we feel very pleased to stay here. The toilet outside the hotel is really clean and the fragrance is really nice. the food that was served is also delicious.

This hotel also located near some of the good restaurants so if you don't know where should you eat, then don't worry
coffee/tea and dessert that was served during my conference is taste really good.

The staff is also really friendly, wherever you meet with the hotel staff, they greet you warmly

There are some notes regarding this hotel:

1. The location actually is really close to Airport, but quite far from the city centre, therefore, it becomes one of the weakness. I think it's good idea to have one near the city centre or Losari beach

2. we booked this hotel via booking. com and we already received the confirmation email, however, when we arrived the receptionist said that they are still not received the confirmation. But they overcome this problem and we could stay there, only have to wait a bit longer time.

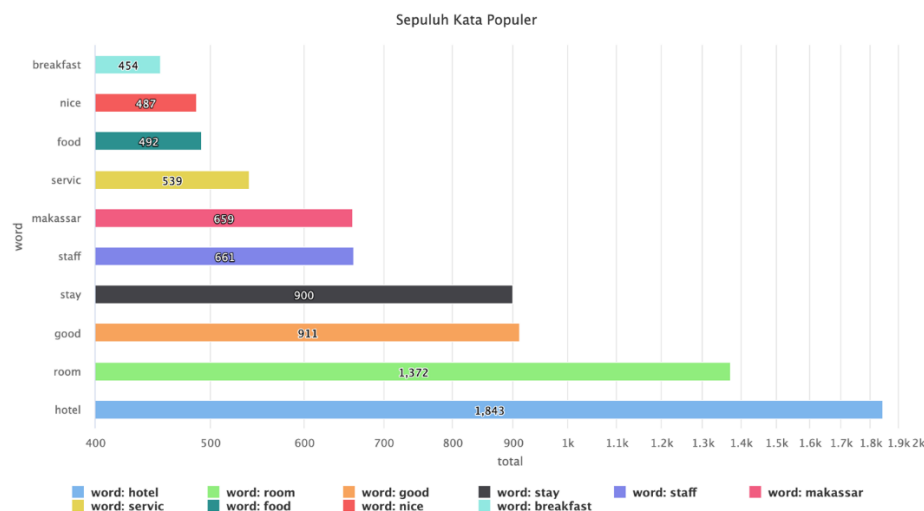
3. We booked 2 non-smoking rooms. When we opened one of the room, it smells like a smoke then the staff provide the stuff to get rid of the smoke smell. I think the hotel should really pay attention and strict about this thing.

4. since this hotel also provides place for gathering in the restaurant, somehow they prepared a BBQ but at that time, when we just came back and goes to the lobby. The BBQ smell was spread off to the lobby. Well, it's no problem since it's the smell of food but also a bit annoyed to smell something like that.

Positif 4,46153846153846

nice (0.46) clean (0.44) pleased (0.49) clean (0.44) nice (0.46) delicious (0.69) good (0.49) worry (-0.49) good (0.49) friendly (0.56) greet (0.33) warmly (0.44) weakness (-0.46) good (0.49) problem (-0.44) like (0.38) pay (-0.10) prepared (0.23) lobby (0.03) lobby (0.03) well (0.28) no (-0.31) problem (-0.44) annoyed (-0.41) like (0.38)

Tabel 1 merupakan hasil ekstrak sentimen dari aplikasi Rapidminer dengan nilai *score* dan *score string* berdasarkan masing-masing data yang telah diproses. Label sentimen negatif dikalkulasi berdasarkan jumlah *scoring string*, sehingga ulasan yang didominasi kata negatif memiliki bobot nilai negatif (-), dan ulasan yang didominasi kata positif memiliki bobot nilai positif (+). Berdasarkan hasil ekstrak sentimen berdasarkan klasifikasi sentimen positif dan negatif, dapat diidentifikasi lima dan sepuluh kata populer yang paling sering muncul dalam klasifikasi negatif dan positif. Dengan demikian, kata-kata populer dalam klasifikasi sentimen negatif dapat digunakan sebagai rekomendasi pengembangan kualitas produk dan layanan hotel, sementara kata-kata populer dalam klasifikasi sentimen positif dapat digunakan sebagai evaluasi untuk mempertahankan kondisi eksisting atau meningkatkan kualitas layanan hotel agar dapat menarik tamu dalam jumlah banyak serta meningkatkan volume penjualan. Berdasarkan hasil identifikasi lima kata populer yang paling sering muncul dalam data ulasan tamu hotel, dapat diketahui bahwa kata hotel, room, good, stay dan staff menjadi kata yang populer. Berdasarkan data tersebut, dapat diinterpretasikan bahwa tamu hotel memiliki pengalaman yang menyenangkan atau memuaskan selama menginap di *top 10 traveler ranked hotel* di Kota Makassar. Selanjutnya, sepuluh kata populer yang paling sering muncul dalam data ulasan beserta jumlah ialah sebagai berikut : hotel (1.843), room (1.372), good (911), stay (900), staff (661), makassar (659), service (539), food (492), nice (487), dan breakfast (454). Adapun, visualisasi sepuluh kata yang paling sering muncul dalam data ulasan *top 10 traveler ranked hotel* di Kota Makassar pada platform Tripadvisor, dapat dilihat pada gambar berikut.



Gambar 3. Sepuluh Kata Populer dari Ulasan Tamu Hotel

Gambar 3 merupakan sepuluh kata populer dari ulasan tamu hotel yang divisualisasikan dalam bentuk *bar-horizontal*, dimana kata hotel muncul sebanyak 1.843, kata room muncul sebanyak 1.372. Hal ini menunjukkan bahwa ekspektasi tamu ketika menggunakan layanan akomodasi hotel merujuk pada jenis dan fasilitas kamar. Dengan demikian, terdapat hubungan antara *room amenities*, *room feature*, dan *room type* dengan preferensi tamu hotel. Selanjutnya, kata *good* muncul sebanyak 911, dan kata *nice* muncul sebanyak 487 dalam data ulasan tamu. Hal ini berarti adanya kesan positif tamu hotel ketika menginap di *top 10 traveler ranked hotel* di Kota Makassar. Adapun, kata *stay* muncul sebanyak 900, dan kata *makassar* muncul sebanyak 659 dalam data ulasan yang mencerminkan kesan tamu hotel ketika menginap di *top 10 traveler ranked hotel* di Kota Makassar yang mana aspek *location* memberikan nilai tambah yang dapat diinterpretasikan lebih jauh pada daya tarik sosial-budaya masyarakat lokal, hingga kondisi ekonomi, politik dan lingkungan. Selanjutnya, kata *staff* muncul sebanyak 661 dan kata *service* muncul sebanyak 539 dalam data ulasan. Berdasarkan hal tersebut dapat diinterpretasikan bahwa pelayanan yang diberikan oleh karyawan di hotel menjadi bagian esensial dalam meningkatkan kenyamanan dan keamanan tamu selama menginap. Adapun, kata *food* muncul sebanyak 492, dan kata *breakfast* muncul sebanyak 454 juga menunjukkan kepuasan tamu terhadap makanan yang disajikan terutama saat sarapan pagi. Dengan demikian, proses ekstraksi dan visualisasi sepuluh kata populer di *top 10 traveler ranked hotel* di Kota Makassar dapat digunakan sebagai rekomendasi untuk mempertahankan maupun meningkatkan produk dan layanan di *top 10 traveler ranked hotel* di Kota Makassar.

Pada tahap pemodelan, data hasil ekstraksi dievaluasi menggunakan algoritma *Decision Tree* (DT) dan algoritma *Support Vector Machine* (SVM) dengan pengaturan data latih sebanyak 30% dan data uji sebanyak (70%) menggunakan operator *split data* pada aplikasi Rapidminer. Proses pengujian performa algoritma dilakukan dengan tahap sebagai berikut : tahap pertama, pengujian tanpa menggunakan operator SMOTE; kedua, pengujian menggunakan operator SMOTE. Adapun, hasil pengujian algoritma *Decision Tree* (DT) dan algoritma *Support Vector Machine* (SVM) tanpa menggunakan operator SMOTE dapat dilihat pada tabel berikut.

Tabel 2. Performa Algoritma DT dan SVM tanpa Operator SMOTE

Decision Tree	Support Vector Machine
PerformanceVector: accuracy: 91.59% +/- 2.82% (micro average: 91.58%) ConfusionMatrix: True: Negatif Positif Negatif: 13 18 Positif: 49 716 AUC (optimistic): 0.984 +/- 0.021 (micro average: 0.984) (positive class: Positif) AUC: 0.591 +/- 0.061 (micro average: 0.591) (positive class: Positif) AUC (pessimistic): 0.199 +/- 0.136 (micro average: 0.199) (positive class: Positif) precision: 93.61% +/- 0.99% (micro average: 93.59%) (positive class: Positif) ConfusionMatrix: True: Negatif Positif Negatif: 13 18 Positif: 49 716 recall: 97.55% +/- 3.58% (micro average: 97.55%) (positive class: Positif) ConfusionMatrix: True: Negatif Positif Negatif: 13 18 Positif: 49 716 f_measure: 95.50% +/- 1.65% (micro average: 95.53%) (positive class: Positif) ConfusionMatrix: True: Negatif Positif Negatif: 13 18 Positif: 49 716	PerformanceVector: accuracy: 92.08% +/- 1.06% (micro average: 92.09%) ConfusionMatrix: True: Negatif Positif Negatif: 1 2 Positif: 61 732 AUC (optimistic): 0.906 +/- 0.058 (micro average: 0.906) (positive class: Positif) AUC: 0.906 +/- 0.058 (micro average: 0.906) (positive class: Positif) AUC (pessimistic): 0.906 +/- 0.058 (micro average: 0.906) (positive class: Positif) precision: 92.31% +/- 0.73% (micro average: 92.31%) (positive class: Positif) ConfusionMatrix: True: Negatif Positif Negatif: 1 2 Positif: 61 732 recall: 99.72% +/- 0.58% (micro average: 99.73%) (positive class: Positif) ConfusionMatrix: True: Negatif Positif Negatif: 1 2 Positif: 61 732 f_measure: 95.87% +/- 0.57% (micro average: 95.87%) (positive class: Positif) ConfusionMatrix: True: Negatif Positif Negatif: 1 2 Positif: 61 732

Tabel 2 merupakan hasil evaluasi performa algoritma *Decision Tree* (DT) dan algoritma *Support Vector Machine* (SVM) tanpa operator SMOTE. Algoritma DT memiliki nilai akurasi 91,59%, nilai presisi 93,61%, nilai *recall* 97,55%, nilai f-measure 95,50%, serta nilai AUC 0,591 (59,1%). Sementara itu algoritma SVM memiliki nilai akurasi 92,08%, nilai presisi 92,31%, nilai *recall* 99,72%, nilai f-measure 95,87%, serta nilai AUC 0,906 (90,6%). Hal ini meenunjukkan bahwa tanpa menggunakan operator SMOTE, performa algoritma SVM lebih baik dalam hal klasifikasi data sentimen dengan pembagian data latih 30% dan data uji 70%. Apabila dilakukan perubahan data latih menjadi 40% dan data uji 60% maka terjadi perubahan nilai akurasi, presisi, *recall*, f-measure, dan nilai AUC pada algoritma DT dan SVM sebagai berikut : pertama, perubahan performa algortima DT ialah nilai akurasi 90,90 %, nilai presisi 93,16, nilai *recall* 97,29%, nilai f-measure 95,17%, dan nilai AUC 56,1%; kedua, perubahan performa algoritma SVM ialah nilai akurasi 99,05%,

nilai presisi 100%, nilai *recall* 98,08%, nilai f-measure 99,02% dan nilai AUC 100%. Hal ini berarti bahwa perubahan pada pembagian jumlah data uji dan data latih juga memengaruhi performa algoritma DT dan SVM. Meskipun demikian, algoritma SVM masih menunjukkan performa terbaik tanpa menggunakan operator SMOTE.

Operator SMOTE digunakan untuk mengatasi persoalan keseimbangan data dimana label sentimen negatif dan positif dalam data set tidak dalam jumlah yang proporsional. Berdasarkan hal tersebut maka evaluasi algoritma DT dan SVM menggunakan operator SMOTE dengan pembagian data latih 40% dan data uji 60% ialah sebagai berikut : pertama, performa algoritma DT menunjukkan nilai akurasi 92,93 %, nilai presisi 94,42, nilai *recall* 91,42%, nilai f-measure 92,64%, dan nilai AUC 88,8%; kedua, performa algoritma SVM menunjukkan nilai akurasi 99,05 %, nilai presisi 100%, nilai *recall* 98,08%, nilai f-measure 99,02%, dan nilai AUC 100%. Hal ini menunjukkan bahwa penggunaan operator SMOTE dapat meningkatkan performa algoritma DT. Meskipun perubahan tersebut tidak berpengaruh signifikan terhadap performa algoritma SVM. Meskipun demikian, apabila pembagian data uji dan data latih diperbaharui menjadi 30% (data latih) dan 70% (data uji) maka dapat diketahui perubahan performa kedua algoritma sebagaimana tabel berikut.

Tabel 3. Performa Algoritma DT dan SVM setelah Menggunakan Operator SMOTE

Decision Tree	Support Vector Machine
PerformanceVector: accuracy: 93.60% +/- 4.48% (micro average: 93.60%)	PerformanceVector: accuracy: 98.98% +/- 0.66% (micro average: 98.98%)
ConfusionMatrix: True: Negatif Positif Negatif: 688 48 Positif: 46 686	ConfusionMatrix: True: Negatif Positif Negatif: 734 15 Positif: 0 719
AUC (optimistic): 0.992 +/- 0.008 (micro average: 0.992) (positive class: Positif)	AUC (optimistic): 1.000 +/- 0.000 (micro average: 1.000) (positive class: Positif)
AUC: 0.901 +/- 0.145 (micro average: 0.901) (positive class: Positif)	AUC: 1.000 +/- 0.000 (micro average: 1.000) (positive class: Positif)
AUC (pessimistic): 0.881 +/- 0.086 (micro average: 0.881) (positive class: Positif)	AUC (pessimistic): 1.000 +/- 0.000 (micro average: 1.000) (positive class: Positif)
precision: 93.92% +/- 2.35% (micro average: 93.72%) (positive class: Positif)	precision: 100.00% +/- 0.00% (micro average: 100.00%) (positive class: Positif)
ConfusionMatrix: True: Negatif Positif Negatif: 688 48 Positif: 46 686	ConfusionMatrix: True: Negatif Positif Negatif: 734 15 Positif: 0 719
recall: 93.47% +/- 10.59% (micro average: 93.46%) (positive class: Positif)	recall: 97.96% +/- 1.32% (micro average: 97.96%) (positive class: Positif)
ConfusionMatrix: True: Negatif Positif Negatif: 688 48 Positif: 46 686	ConfusionMatrix: True: Negatif Positif Negatif: 734 15 Positif: 0 719
f_measure: 93.31% +/- 5.33% (micro average: 93.59%) (positive class: Positif)	f_measure: 98.97% +/- 0.68% (micro average: 98.97%) (positive class: Positif)
ConfusionMatrix: True: Negatif Positif Negatif: 688 48 Positif: 46 686	ConfusionMatrix: True: Negatif Positif Negatif: 734 15 Positif: 0 719

Tabel 3 merupakan hasil evaluasi performa algoritma *Decision Tree* (DT) dan algoritma *Support Vector Machine* (SVM) setelah menggunakan operator SMOTE dengan jumlah data latih 30% dan data uji 70%. Secara spesifik, performa algoritma DT menunjukkan nilai akurasi 93,60%, nilai presisi 93,92%, nilai *recall* 93,47%, nilai f-measure 93,31% dan nilai AUC sebesar 0,901 (90,1%). Sementara itu, performa algoritma SVM menunjukkan nilai akurasi 98,98%, nilai presisi 100%, nilai *recall* 97,96%, nilai f-measure 98,97%, dan nilai AUC sebesar 1.000 (100%). Hal ini menunjukkan bahwa algoritma DT dan SVM ketika menggunakan operator SMOTE dengan pembagian data latih sebanyak 30% dan data uji 70% masih menunjukkan performa yang baik. Selanjutnya, apabila jumlah data latih dikurangi menjadi 20% dan jumlah data uji ditingkatkan sebesar 80% maka dapat diketahui bahwa algoritma DT menunjukkan nilai akurasi 96,42%, nilai presisi 94,85%, nilai *recall* 98,21%, nilai f-measure 96,49%, dan nilai AUC sebesar 0,963 (96,3%). Sedangkan algoritma SVM menunjukkan nilai akurasi 99,05%, nilai presisi 100%, nilai *recall* 98,10%, nilai f-measure 99,03%, dan nilai AUC sebesar 1.000 (100%). Hal ini berarti bahwa setelah perubahan jumlah data uji dan data latih, algoritma SVM menunjukkan performa yang lebih baik dibandingkan dengan DT. Mempertimbangkan hal tersebut maka penelitian ini merekomendasikan algoritma SVM sebagai model yang relevan dengan konteks data teks untuk analisis sentimen *top 10 traveler ranked hotel* di Kota Makassar.

Penelitian terdahulu menunjukkan bahwa algoritma SVM dalam beberapa studi kasus menunjukkan performa yang lebih baik dibandingkan algoritma lainnya. Novichasari dan Rakhmawati menunjukkan bahwa performa algoritma *Support Vector Machine* (SVM) jauh lebih unggul dibandingkan *Neural Network* (NN) dalam analisis data time-series dan non-time series [20]. Disisi lain, Nada et al. menunjukkan bahwa performa algoritma *Support Vector Machine* (SVM) jauh lebih unggul dibandingkan *Naïve Bayes Classifier* (NBC) dalam klasifikasi data sentimen berbasis media sosial

Twitter menanggapi isu BPJS di Indonesia [21]. Adapun, Gumelar et al. mengemukakan bahwa keseimbangan data dalam proses klasifikasi berperan penting dalam menentukan performa algoritma, oleh sebab itu operator SMOTE dapat digunakan untuk mengatasi data teks dengan label yang tidak seimbang [22]. Hal ini berarti bahwa studi tentang analisis sentimen yang membandingkan performa SVM dengan algoritma lain, menunjukkan keunggulan algoritma SVM dengan menggunakan operator SMOTE maupun tanpa menggunakan operator SMOTE. Meskipun demikian, pengujian algoritma merupakan tahap modeling dalam metode CRISP-DM, dimana tahap *business understanding*, tahap *data understanding*, dan tahap *data preparation* berperan penting dalam menunjukkan konteks penambahan dan pengolahan data teks menjadi informasi yang actual dan dapat digunakan untuk pengembangan bisnis atau organisasi.

4. KESIMPULAN

Hasil penelitian ini menunjukkan bahwa implementasi metode CRISP-DM dalam identifikasi dan analisis sentimen tamu terhadap produk dan layanan di *10 traveler ranked hotel* di Kota Makassar versi Tripadvisor menunjukkan bahwa terdapat sepuluh kata populer yang paling sering muncul dalam data ulasan setelah proses ekstraksi, kata-kata tersebut berhubungan dengan aspek *location*, *cleanliness*, *service*, dan *value*, secara spesifik juga berhubungan dengan *room amenity*, *room feature*, dan *room type* yang dimiliki masing-masing hotel. Adapun kata-kata populer yang dimaksud ialah sebagai berikut : hotel (1.843), room (1.372), good (911), stay (900), staff (661), makassar (659), service (539), food (492), nice (487), dan breakfast (487). Selain itu, hasil perbandingan algoritma DT dan SVM menunjukkan bahwa algoritma dengan performa terbaik ialah SVM, sehingga dapat dijadikan model klasifikasi data sentimen dalam penelitian ini. Adapun, berdasarkan pembagian data latih sebanyak 30% dan data uji 70% diperoleh nilai akurasi 98,98%, nilai presisi 100%, nilai recall 97,96%, nilai *f-measure* 98,97%, dan nilai AUC sebesar 1.000 (100%).

UCAPAN TERIMA KASIH

Terima kasih disampaikan kepada pihak-pihak yang telah mendukung terlaksananya penelitian ini, peneliti mengucapkan terima kasih kepada Program Studi Pariwisata, Program Studi Sistem Informasi, Fakultas Teknik, Fakultas Ilmu Administrasi Bisnis dan Ilmu Komunikasi, Lembaga Penelitian dan Pengabdian kepada Masyarakat, Universitas Katolik Indonesia Atma Jaya.

REFERENCES

- [1] H. Yuliamir, Y. S. R, S. Hadi, E. Rahayu, and D. Palupiningtyas, "Peranan kepuasan dalam memediasi pengaruh kualitas layanan , customer experince , dan kepercayaan terhadap loyalitas konsumen hotel bintang 4 di Kota Semarang," *J. Manaj. Perhotelan dan Pariwisata*, vol. 6, no. 1, pp. 42–53, 2023.
- [2] V. L. C. Runtuwarouw, T. O. Rotinsulu, and ..., "Analisis Pengaruh Arus Kunjungan Wisata, Jumlah Hotel Dan Restoran Terhadap Pertumbuhan Ekonomi Di Kota Tomohon Periode ...," *J. Berk. Ilm. ...*, vol. 22, no. 7, pp. 13–24, 2022, [Online]. Available: <https://ejournal.unsrat.ac.id/index.php/jbie/article/view/44251%0Ahttps://ejournal.unsrat.ac.id/index.php/jbie/article/download/44251/38539>
- [3] Y. S. Halum, E. H. Selamat, F. Jemadi, and S. Hame, "Pendampingan Masyarakat Desa Wisata Sano Nggoang Sebagai Langkah Awal Penetapan Destinasi Wisata Super Prioritas Komodo-Labuan Bajo," *J. Pengabd. Kpd. Masy.*, vol. 27, no. 4, pp. 308–313, 2021.
- [4] R. Kurniansah and J. J. Purnama, "Komponen-Komponen Pendukung Pariwisata Kuta Lombok Provinsi Nusa Tenggara Barat," *J. Hosp. dan Pariwisata*, vol. 6, no. 1, pp. 1–6, 2020, doi: 10.30813/jhp.v6i1.2029.
- [5] A. Dianindarwati, B. Setiawati, and J. Usman, "Implementasi Kebijakan Pengembangan Obyek Wisata Pantai Pada Dinas Kebudayaan dan Pariwisata di Kota Makassar," *Kaji. Ilm. Mhs. Adm. Publik*, vol. 3, no. 6, pp. 1662–1672, 2022.
- [6] Hardianti and Nasrullah, "Pengaruh Bauran Promosi Terhadap Keputusan Wisatawan Berkunjung pada Fort Rotterdam di Kota Makasar," *J. Pariwisata Budaya dan Keagamaan*, vol. 1, no. 2, pp. 84–91, 2023.
- [7] A. S. Sasmita, R. Hayati, and P. Sasmita, "Implementasi Prosedur Kerja Guest Service Agent Terhadap Penerapan Adaptasi Kebiasaan Baru Dalam Melayani Tamu Check-in Dan Check-Out Di 5 Hotel Bintang 4 Kota Makassar," *Hosp. Gastron. Res. J.*, vol. 5, no. 1, pp. 48–63, 2023.
- [8] M. Fiqih, O. Hattah, M. Asdar, and I. Nursyamsi, "Analisis Pengaruh Bauran Pemasaran terhadap Daya Tarik Pelanggan pada Hotel Claro di Kota Makassar," *SEIKO J. Manag. Bus.*, vol. 6, no. 1, pp. 208–224, 2023, doi: 10.37531/sejaman.v6i1.3845.
- [9] I. G. A. P. Arimbawa, A. Goeliling, T. Ripal, A. B. Peringga Putra, and A. R. Amin, "The Effect of Communication, Motivation and Job Satisfaction on Employee Performance: Empirical Study Of Hotel Harper Prints By Aston, Makassar," *Pusaka J. Tour. Hosp. Travel Bus. Event*, vol. 5, no. 1, pp. 20–32, 2023, doi: 10.33649/pusaka.v5i1.195.
- [10] T. Clarita, "Analisis Sentimen Emosional pada Tamu Hotel yang Menginap di Wilayah Isolasi Mandiri Covid-19 (Analisis Konten pada Tamu Hotel Melalui Laman Web Tripadvisor)," *SABA J. Tour. Res.*, vol. 1, no. 1, pp. 32–44, 2023.
- [11] Y. Harwani, "Constructing Brand Personality from the TripAdvisor Online Reviews," *Int. J. Bus. Manag. Technol.*,

- vol. 5, no. 4, pp. 152–157, 2021, [Online]. Available: <https://www.theijbmt.com/archive/0940/1938740165.pdf>
- [12] Y. L. Sukmana, “Implementasi CRISP-DM dengan Menggunakan Tableau untuk Visualisasi Data dalam Upaya Peningkatan Impor Pada KPU Bea Cukai Tipe B Batam,” *J. Perspekt. Bea dan Cukai*, vol. 7, no. 1, pp. 19–38, 2023.
- [13] Y. Yudianta, A. Y. Agustina, and N. Khofifah, “Prediksi Customer Churn Menggunakan Metode CRISP-DM Pada Industri Telekomunikasi Sebagai Implementasi Mempertahankan Pelanggan,” *IJIEB Indones. J. Islam. Econ. Bus.*, vol. 8, no. 1, pp. 1–20, 2023, [Online]. Available: <http://e-journal.lp2m.uinjambi.ac.id/ojp/index.php/ijoieb>
- [14] L. A. Pramesti and N. Pratiwi, “Analisis Sentimen Twitter Terhadap Program MBKM Menggunakan Decision Tree dan Support Vector Machine,” *J. Inf. Syst. Res.*, vol. 4, no. 4, pp. 1145–1154, 2023, doi: 10.47065/josh.v4i4.3807.
- [15] N. T. Romadloni and W. Supriyanti, “Analisis Sentimen Penggunaan Teknologi Pada Pendidikan Anak Usia Dini,” *J. Ilm. Sinus*, vol. 21, no. 2, pp. 101–109, 2023.
- [16] M. F. Asshiddiqi and K. M. Lhaksmana, “Perbandingan Metode Decision Tree dan Support Vector Machine untuk Analisis Sentimen pada Instagram Mengenai Kinerja PSSI,” in *e-Proceeding of Engineering*, 2020, vol. 7, no. 3, pp. 9936–9948.
- [17] D. N. Fitriana and Y. Sibaroni, “Sentiment Analysis on KAI Twitter Post Using Multiclass Support Vector Machine (SVM),” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 5, pp. 846–853, 2020, doi: 10.29207/resti.v4i5.2231.
- [18] A. Karim, “Perbandingan Prediksi Kemiskinan di Indonesia Menggunakan Support Vector Machine (SVM) dengan Regresi Linear,” *J. Sains Mat. dan Stat.*, vol. 6, no. 1, pp. 107–112, 2020, doi: 10.24014/jsms.v6i1.9259.
- [19] E. A. Nida, “Analisis Kinerja Algoritma Support Vector Machine (SVM) Guna Pengambilan Keputusan Beli/Jual Pada Saham PT Elnusa Tbk. (ELSA),” *J. Transform.*, vol. 17, no. 2, pp. 160–170, 2020, doi: 10.26623/transformatika.v17i2.1649.
- [20] S. I. Novichasari and R. Rakhmawati, “Komparasi Neural Network dan Support Vector Machine untuk Data Time Series dan Non-Time Series,” *J. Prodi Tek. Inform. UNW*, vol. 5, no. 1, pp. 4–9, 2023.
- [21] D. D. Nada, Soehardjoepri, and R. M. Atok, “Perbandingan Analisis Sentimen Mengenai BPJS pada Media Sosial Twitter Menggunakan Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM),” *J. Sains dan Seni ITS*, vol. 11, no. 6, p. 480=485, 2023, doi: 10.12962/j23373520.v11i6.96330.
- [22] G. Gumelar, Kusrini, and H. Al Fatta, “Kombinasi Algoritma Klasifikasi Dengan Algoritma Oversampling Untuk Menangani Ketidakseimbangan Kelas Pada Level Data,” *J. Tek. Inform. dan Sist. Inf.*, vol. 10, no. 2, pp. 29–39, 2023.