

KOMPARASI ALGORITMA KLASIFIKASI PADA ANALISIS REVIEW HOTEL

Lila Dini Utami¹; Hilda Rachmi²; Dini Nurlaela³

Sistem Informasi Akuntansi
Universitas Bina Sarana Informatika
www.bsi.ac.id
lila.ldu@bsi.ac.id

Sistem Informasi Kampus Kota Bogor
Universitas Bina Sarana Informatika
www.bsi.ac.id
hilda.hlr@bsi.ac.id

Ilmu Komputer
STMIK Nusa Mandiri Jakarta
www.nusamandiri.ac.id
dinidur.dur@gmail.com



Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi-NonKomersial 4.0 Internasional.

Abstract - *At this time the freedom to express opinions in oral and written forms about everything is very easy. This activity can be used to make decisions by some business people. Especially by service providers, such as hotels. This will be very useful in the development of the hotel business itself. But the review data must be processed using the right algorithm. So this study was conducted to find out which algorithms are more feasible to use to get the highest accuracy. The methods used are Naïve Bayes (NB), Support Vector Machine (SVM), and k-Nearest Neighbor (k-NN). From the process that has been done, the results of Naïve Bayes accuracy are 71.50% with the AUC value is 0.500, Support Vector Machine is 72.50% with the AUC value is 0.936 and the accuracy results if using the k-Nearest Neighbor algorithm is 75.00% with the AUC value is 0.500. The use of the k-Nearest Neighbor algorithm can help in making more appropriate decisions for hotel reviews at this time.*

Intisari – Pada saat ini kebebasan menyampaikan pendapat dalam bentuk lisan maupun tulisan tentang segala hal sudah sangat mudah. Kegiatan ini dapat dimanfaatkan untuk mengambil keputusan oleh sebagian pelaku bisnis. Khususnya oleh penyedia jasa, seperti hotel. Hal itu akan sangat bermanfaat dalam perkembangan bisnis hotel itu sendiri. Namun data review tersebut harus diolah menggunakan algoritma yang tepat. Maka penelitian ini dilakukan untuk mengetahui algoritma yang lebih layak digunakan untuk mendapatkan akurasi yang paling tinggi. Adapun

metode yang digunakan adalah Naïve Bayes (NB), Support Vector Machine (SVM), dan k-Nearest Neighbor (k-NN). Dari proses yang telah dilakukan didapatkan hasil akurasi Naïve Bayes adalah 71,50% dengan nilai AUC adalah 0,500, Support Vector Machine adalah 72,50% dengan nilai AUC adalah 0,936 dan hasil akurasi jika menggunakan algoritma k-Nearest Neighbor adalah 75,00% dengan nilai AUC adalah 0,500. Penggunaan algoritma k-Nearest Neighbor dapat membantu dalam pengambilan keputusan yang lebih tepat untuk review hotel pada saat ini.

Kata Kunci: *naïve bayes, support vector machine, k-nn, text mining, review hotel, komparasi, algoritma klasifikasi.*

PENDAHULUAN

Penggunaan internet sebagai sumber pencarian data menjadi hal yang tidak asing lagi bagi hampir seluruh lapisan masyarakat. Peningkatan pengguna internet semakin melonjak dari tahun ke tahun. Pencarian data yang dilakukan bukan hanya untuk memenuhi kebutuhan pokok, namun juga sebagai alat untuk pemenuhan kebutuhan tersier seperti berwisata.

Rencana liburan dan pemesanan hotel melalui situs menjadi salah satu penggunaan komersial yang dianggap penting. Berbagi melalui situs telah menjadi alat utama untuk mengungkapkan pendapat pelanggan mengenai barang atau layanan tertentu (Elango & Narayan, 2011). Hal ini

menunjukkan bahwasanya saat ini sebelum memutuskan untuk melakukan pemesanan hotel ataupun melakukan perjalanan wisata, pelanggan akan terlebih dahulu mencari review atau pendapat orang lain yang sebelumnya telah menggunakan hotel tersebut.

Berbagai macam situs wisata yang menyediakan layanan untuk memudahkan wisatawan semakin mudah dijumpai, salah satunya adalah situs www.tripadvisor.com. Situs ini menyediakan fitur pencarian hotel yang dilengkapi dengan review serta pemberian tingkatan bintang mulai dari bintang 1 hingga bintang 5 yang diberikan oleh pengguna. Review yang dituliskan pun bermacam-macam, mulai dari keadaan fasilitas yang disediakan, lokasi hotel dengan pusat perbelanjaan atau tempat wisata lain serta kenyamanan dan layanan yang diberikan oleh pihak hotel.

Namun peringkat skalar dengan skor antara 1 – 5 sangat tidak membantu manajer hotel karena nilai numerik tidak memberikan informasi mengenai apa yang sebenarnya diinginkan atau dianggap positif oleh tamu hotel. Juga nilai numerik tersebut tidak dapat dibandingkan, ketika sebuah hotel bintang tiga mendapat skor yang lebih tinggi dari skor bintang empat tidak menyiratkan hotel yang satu lebih baik dari yang lain (Sodanil, 2016). Di sisi lain penelitian yang dilakukan di Cina dengan mengumpulkan 58 hotel yang memiliki rating 3-5 berdasarkan indeks popularitas hotel mengidentifikasi sejumlah kategori dari review yang diberikan secara online, sentimen dari masing-masing kategori dan hubungan antar kategori. Penelitian ini menemukan tingkatan bintang yang diberikan berkorelasi baik dengan skor sentimen untuk judul maupun isi konten review (Tian, Tao, He, & Akula, 2016).

Pada penelitian yang dilakukan oleh (Yordanova & Kabakchieva, 2017) mengungkapkan bahwa tantangan awal saat melakukan analisis terhadap respon dari tamu hotel adalah memperkirakan pendapat yang dinyatakan dalam review hotel dengan mengklasifikasikan respon tersebut sebagai umpan balik positif atau negatif. Hal ini bisa diatasi dengan mengaplikasikan sentimen analisis. Tantangan kedua adalah membayangkan hasil dengan tujuan untuk mengekstrak pengetahuan bisnis yang didapatkan melalui alat Business Intelligence. Sentimen analisis menurut (Liu, 2012) atau yang disebut juga sebagai *opinion mining* adalah bidang studi yang menganalisis opini orang-orang, sentimen, evaluasi, penilaian, sikap dan emosi terhadap entitas seperti produk, layanan, organisasi, perorangan, masalah, peristiwa, topik dan berbagai atributnya.

Sentimen analisis (Sipayung, Maharani, & Zefanya, 2016) adalah metode untuk memahami, mengekstrak, dan mengolah data ulasan atau tekstual untuk mendapatkan sentimen yang terkandung dalam sebuah ulasan. Ulasan dapat dibagi menjadi 3 jenis, yaitu positif, negatif dan netral. Dengan menggunakan sentimen analisis pihak perusahaan dapat mengetahui respon pelanggan terhadap suatu produk atau pelayanan melalui feedback yang diberikan.

Beberapa tahun terakhir, telah banyak dilakukan penelitian terkait dengan analisis review, diantaranya:

1. Pada penelitian ulasan film dengan fokus mengklasifikasikan ulasan kedalam dua kelas, yaitu positif dan negatif menggunakan algoritma Support Vector Machine (SVM), akurasi yang didapat sebesar 71,87% dan meningkat menjadi 77% dengan menggunakan hybrid Particle Swarm Optimization (PSO) (Basari, Hussin, Ananta, & Zeniarja, 2013)
2. *Opinion mining* pada review hotel menggunakan algoritma Naive-Bayes yang dilakukan oleh (Muthia, 2015). Penelitian ini menunjukkan pengaruh penggunaan fitur n-gram pada Naive Bayes dengan peningkatan akurasi menjadi 87%.
3. Peneliti lain membandingkan algoritma klasifikasi antara Support Vector Machine (SVM), Naive Bayes (NB), dan Artificial Neural Network untuk ulasan film dengan akurasi tertinggi dihasilkan oleh SVM dengan nilai akurasi 81,10%. Sedangkan hasil dari komparasi algoritma *feature selection* antara *information gain*, *chi square*, *forward selection*, dan *backward elimination* didapatkan hasil tertinggi pada *information gain* menggunakan parameter top k dengan nilai k = 200 dan hasil akurasi 84,57% (Chandani, 2015).
4. Penelitian yang dilakukan (Dey, Chakraborty, Biswas, Bose, & Tiwari, 2016) membahas dua algoritma *machine learning*, yaitu K-Nearest Neighbor (K-NN) dan Naive Bayes serta membandingkan kedua akurasi secara keseluruhan. Hasilnya menunjukkan dalam ulasan film, Naive Bayes memberikan hasil yang jauh lebih baik dari K-NN tetapi pada ulasan hotel keduanya memberikan performa yang hampir sama.
5. Penelitian terkait dengan klasifikasi sentimen mengenai review hotel di sosial media menggunakan beberapa algoritma seperti Decision Tree Learning, K-NN, dan Neural Networks. Dengan menggunakan dataset yang seimbang 20-200, didapatkan akurasi tertinggi sebesar 85.90% (Yordanova & Kabakchieva, 2017)

Meskipun jumlah penelitian tentang aplikasi *text mining* mengalami peningkatan, dalam pemasaran, namun masih ada kesenjangan penelitian mengenai preferensi penilaian pelanggan melalui prosedur *text mining* (Dirsehan, 2016). Sampai saat ini telah banyak metode yang digunakan untuk mendapatkan akurasi tertinggi dalam penelitian bidang *text mining*. Penambahan fitur ke dalam beberapa metode juga telah dilakukan. Namun tetap saja hasil yang berbeda ditunjukkan oleh penelitian yang berbeda. Dan belum ada kesimpulan tetap yang bisa membuktikan algoritma terbaik untuk penelitian yang melibatkan sentimen analisis. Oleh karena itu penelitian ini bertujuan untuk mendapatkan algoritma terbaik antara beberapa algoritma yang dapat digunakan dalam pengujian analisis sentimen terkait dengan review hotel. Agar pengujian analisis sentimen, khususnya yang dilakukan pada salah satu situs pencarian hotel yang banyak digunakan saat ini bisa mendapatkan hasil yang optimal.

BAHAN DAN METODE

Peneliti melakukan komparasi tiga algoritma, yaitu Naïve Bayes (NB), Support Vector Machine (SVM), dan k-Nearest Neighbor (k-NN). Penelitian ini bertujuan untuk mengetahui akurasi dari masing-masing algoritma.

Adapun tahapan penelitian yang dilakukan oleh peneliti, sebagai berikut :

1. Pengumpulan data

Penulis menggunakan data review salah satu hotel yang berada di Hawaii, yang diambil dari situ www.tripadvisor.com. Review hotel yang digunakan terdiri dari 100 review positif dan 100 review negatif dengan teks berbahasa Inggris.

2. Pengolahan awal data

Dalam tahapan ini, terdapat dua proses, yaitu:

a. Tokenization

Dalam proses ini, semua kata yang ada didalam tiap dokumen dikumpulkan dan dihilangkan tanda bacanya, serta dihilangkan jika terdapat simbol atau apapun yang bukan huruf.

Tabel 1. Hasil Proses Tokenization

Review	Tokenization
I stayed at the Andaz at Wailea with my wife in January. The resort is nice, but not worth the price. The good: It comes with a lot of amenities included--guided kayak tour, outrigger canoe ride, stand-up-paddle board tour, snorkel gear, boogie boards, beach chairs and	I stayed at the Andaz at Wailea with my wife in January The resort is nice but not worth the price The good It comes with a lot of amenities included guided kayak tour outrigger canoe ride stand up paddle board tour snorkel gear boogie boards beach chairs and

umbrellas, and some great pools. The room was top notch, including the bathroom and beds were comfortable as well. The badThe food was good (and expectedly overpriced) but dinner at the Ka'an restaurant was served by inattentive staff and took far too long. The beach stuff can't go anywhere. You can't just use one of the included kayaks, you have to go out with a "guide;" same with the Stand Up Paddle boards. There's nothing but resorts around here. Further up the coast in Lahaina, you get very similar beaches plus a great town with plenty of eating options and some shopping. The only thing around here is a second rate mall with a Tiffany's, Gap, and a couple of lunch places similar to a Red Robin. Maui was amazing, but you have to leave Wailea to really enjoy it, and the niceness of the resort didn't make up for the lack of food options.

umbrellas and some great pools The room was top notch including the bathroom and beds were comfortable as well The badThe food was good and expectedly overpriced but dinner at the Ka an restaurant was served by inattentive staff and took far too long The beach stuff can t go anywhere You can t just use one of the included kayaks you have to go out with a guide same with the Stand Up Paddle boards There s nothing but resorts around here Further up the coast in Lahaina you get very similar beaches plus a great town with plenty of eating options and some shopping The only thing around here is a second rate mall with a Tiffany s Gap and a couple of lunch places similar to a Red Robin Maui was amazing but you have to leave Wailea to really enjoy it and the niceness of the resort didn t make up for the lack of food options

Sumber : (Utami, Rachmi, & Nurlaela, 2018)

b. Stopwords Removal

Yaitu penghapusan kata-kata yang tidak relevan, seperti the, of, for, with, dan sebagainya

Tabel 2. Hasil Proses Stopwords Removal

Review	Stopwords Removal
I stayed at the Andaz at Wailea with my wife in January. The resort is nice, but not worth the price. The good: It comes with a lot of amenities included--guided kayak tour, outrigger canoe ride, stand-up-paddle board tour, snorkel gear, boogie boards, beach chairs and umbrellas, and some great pools. The room was top notch, including the bathroom and beds were comfortable as well. The badThe food was good (and expectedly overpriced) but dinner at the Ka'an restaurant was served by inattentive staff and took far too long. The beach stuff can't go anywhere. You can't just use one of the included kayaks, you have to go out with a "guide;" same with the Stand Up Paddle boards. There's nothing but resorts around here. Further up the coast in Lahaina, you get	I stayed Andaz Wailea wife January resort nice worth price good comes lot amenities included guided kayak tour outrigger canoe ride stand paddle board tour snorkel gear boogie boards beach chairs umbrellas great pools room top notch including bathroom beds comfortable badThe food good expectedly overpriced dinner Ka restaurant served inattentive staff took beach stuff t go anywhere t use included kayaks go guide Stand Paddle boards s resorts Further coast Lahaina get similar beaches great town plenty eating options shopping thing rate mall Tiffany s Gap couple lunch places similar Red Robin Maui amazing leave Wailea enjoy niceness resort t make lack food options

very similar beaches plus a great town with plenty of eating options and some shopping. The only thing around here is a second rate mall with a Tiffany's, Gap, and a couple of lunch places similar to a Red Robin. Maui was amazing, but you have to leave Wailea to really enjoy it, and the niceness of the resort didn't make up for the lack of food options.

Sumber : (Utami, Rachmi, & Nurlaela, 2018)

c. Stemming

Yaitu mengelompokkan kata kedalam beberapa kelompok yang memiliki kata dasar yang sama, seperti drug, drugged, dan drugs dimana kata dasar dari semuanya adalah kata drug.

Tabel 3. Hasil Proses Stemming

Review	Stopwords Removal
I stayed at the Andaz at Wailea with my wife in January. The resort is nice, but not worth the price. The good: It comes with a lot of amenities included-guided kayak tour, outrigger canoe ride, stand-up-paddle board tour, snorkel gear, boogie boards, beach chairs and umbrellas, and some great pools. The room was top notch, including the bathroom and beds were comfortable as well. The badThe food was good (and expectedly overpriced) but dinner at the Ka'an restaurant was served by inattentive staff and took far too long. The beach stuff can't go anywhere. You can't just use one of the included kayaks, you have to go out with a "guide;" same with the Stand Up Paddle boards. There's nothing but resorts around here. Further up the coast in Lahaina, you get very similar beaches plus a great town with plenty of eating options and some shopping. The only thing around here is a second rate mall with a Tiffany's, Gap, and a couple of lunch places similar to a Red Robin. Maui was amazing, but you have to leave Wailea to really enjoy it, and the niceness of the resort didn't make up for the lack of food options.	i stay andaz wailea wife januari resort nice worth price good come lot amen includ guid kayak tour outrigg cano ride stand paddl board tour snorkel gear boogi board beach chair umbrella great pool room top notch includ bathroom bed comfort badth food good expect overpr dinner ka restaur serv inattent staff took beach stuff t go anywher t use includ kayak go guid stand paddl board s resort further coast lahaina get similar beach great town plenti eat option shop thing rate mall tiffani s gap coupl lunch place similar red robin maui amaz leav wailea enjoy nice resort t make lack food option

Sumber : (Utami, Rachmi, & Nurlaela, 2018)

3. Metode Yang Diusulkan

Metode yang peneliti usulkan adalah mengenai klasifikasi teks review hotel dengan menggunakan Algoritma Naïve Bayes (NB), Support Vector Machine (SVM), dan k-Nearest Neighbor (k-NN) sebagai perbandingan akurasi. Dataset yang digunakan berasal dari www.tripadvisor.com yang terdiri dari 100 review positif dan 100 review negatif.

Untuk *preprocessing* dilakukan *tokenize*, *stopwords removal*, dan *stemming*. Penelitian ini nantinya menghasilkan akurasi dan nilai AUC dan menggunakan RapidMiner versi 5.3

HASIL DAN PEMBAHASAN

1. Hasil Menggunakan Algoritma Naïve Bayes

Dari 200 data review hotel yaitu 100 review positif dan 100 review negatif, sebanyak 93 data diprediksi sesuai yaitu negatif dan sebanyak 7 data diprediksi negatif tetapi ternyata positif. Sebaliknya, 50 data diprediksi sesuai yaitu positif dan sebanyak 50 data diprediksi positif tetapi ternyata negatif.

Hasil akurasi jika menggunakan algoritma Naïve Bayes adalah 71,50% dengan nilai AUC adalah 0,500.

Tabel 1. Confusion Matrix Algoritma Naïve Bayes

Accuracy: 71,50% +/- 6,73% (mikro: 71,50%)			
	True Negatif	True Positif	Class Precision
Pred. Negatif	93	50	65,03%
Pred. Positif	7	50	87,72%
Class Recall	93,00%	50,00%	

Sumber : (Utami, Rachmi, & Nurlaela, 2018)

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

$$Accuracy = \frac{(93 + 50)}{(93 + 7 + 50 + 50)}$$

$$Accuracy = \frac{143}{200} = 0,715 = 71,50\%$$



Sumber : (Utami, Rachmi, & Nurlaela, 2018)

Gambar 1. Grafik Area Under Curve (AUC) Algoritma Naïve Bayes

2. Hasil Menggunakan Algoritma Support Vector Machine

Dari 200 data review hotel yaitu 100 review positif dan 100 review negatif, sebanyak 47 data diprediksi sesuai yaitu negatif dan sebanyak 53 data diprediksi negatif tetapi ternyata positif. Sebaliknya, 98 data diprediksi sesuai yaitu positif dan sebanyak 2 data diprediksi positif tetapi ternyata negatif.

Hasil akurasi jika menggunakan algoritma Support Vector Machine adalah 72,50% dengan nilai AUC adalah 0,936.

Tabel 2. Confusion Matrix Algoritma Support Vector Machine

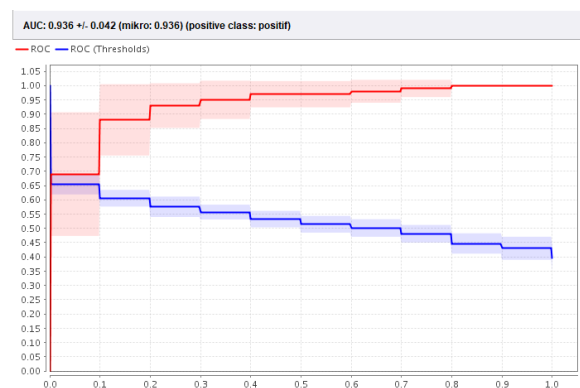
Accuracy: 72,50% +/- 9,55% (mikro: 72,50%)			
	True Negatif	True Positif	Class Precision
Pred. Negatif	47	2	95,92%
Pred. Positif	53	98	64,90%
Class Recall	47,00%	98,00%	

Sumber : (Utami, Rachmi, & Nurlaela, 2018)

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

$$Accuracy = \frac{(47 + 98)}{(47 + 53 + 98 + 2)}$$

$$Accuracy = \frac{145}{200} = 0,725 = 72,50\%$$



Gambar 2. Grafik Area Under Curve (AUC) Algoritma Support Vector Machine

3. Hasil Menggunakan Algoritma k-Nearest Neighbor

Dari 200 data review hotel yaitu 100 review positif dan 100 review negatif, sebanyak 47 data diprediksi sesuai yaitu negatif dan sebanyak 53 data diprediksi negatif tetapi ternyata positif. Sebaliknya, 98 data diprediksi sesuai yaitu positif dan sebanyak 2 data diprediksi positif tetapi ternyata negatif.

Sebaliknya, 98 data diprediksi sesuai yaitu positif dan sebanyak 2 data diprediksi positif tetapi ternyata negatif.

Hasil akurasi jika menggunakan algoritma k-Nearest Neighbor adalah 75,00% dengan nilai AUC adalah 0,500.

Tabel 3. Confusion Matrix Algoritma k-Nearest Neighbor

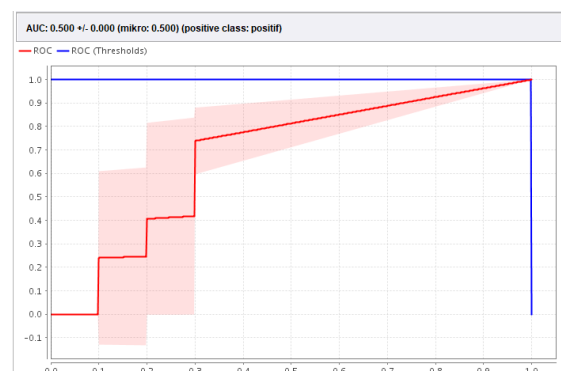
Accuracy: 75,00% +/- 9,75% (mikro: 75,00%)			
	True Negatif	True Positif	Class Precision
Pred. Negatif	78	28	73,58%
Pred. Positif	22	72	76,60%
Class Recall	78,00%	72,00%	

Sumber : (Utami, Rachmi, & Nurlaela, 2018)

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

$$Accuracy = \frac{(78 + 72)}{(78 + 22 + 72 + 28)}$$

$$Accuracy = \frac{150}{200} = 0,75 = 75,00\%$$



Sumber : (Utami, Rachmi, & Nurlaela, 2018)
Gambar 3. Grafik Area Under Curve (AUC) Algoritma k-Nearest Neighbor

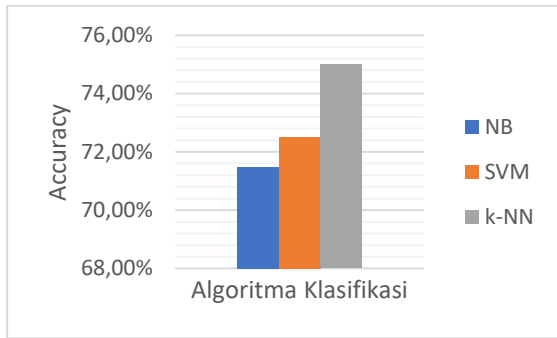
4. Hasil Penelitian

Rangkuman hasil ketiga komparasi algoritma klasifikasi tersebut, sebagai berikut :

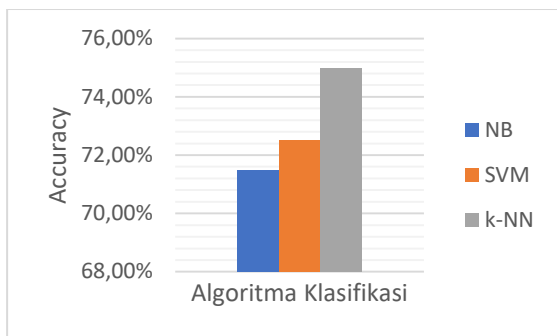
Tabel 4. Komparasi Accuracy dan AUC Algoritma Klasifikasi

	Accuracy	AUC
NB	71,50%	0,500
SVM	72,50%	0,936
k-NN	75,00%	0,500

Sumber : (Utami, Rachmi, & Nurlaela, 2018)



Sumber : (Utami, Rachmi, & Nurlaela, 2018)
 Gambar 4. Komparasi Accuracy Algoritma Klasifikasi



Sumber : (Utami, Rachmi, & Nurlaela, 2018)
 Gambar 5. Komparasi AUC Algoritma Klasifikasi

KESIMPULAN

Hasil dari komparasi algoritma klasifikasi antara Naïve Bayes (NB), Support Vector Machine (SVM), dan k-Nearest Neighbor (k-NN) didapatkan k-NN dengan hasil terbaik dengan akurasi 75,00% dengan AUC adalah 0,500.

REFERENSI

Basari, A. S. H., Hussin, B., Ananta, I. G. P., & Zeniarja, J. (2013). Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization. *Procedia Engineering*, 53(September 2015), 453-462. <https://doi.org/10.1016/j.proeng.2013.02.059>

Chandani, V. (2015). Komparasi Algoritma Klasifikasi Machine Learning Dan Feature Selection pada Analisis Sentimen Review Film, 1(1), 55-59.

Dey, L., Chakraborty, S., Biswas, A., Bose, B., & Tiwari, S. (2016). Sentiment Analysis of Review Datasets Using Naïve Bayes' and K-NN Classifier. *International Journal of Information Engineering and Electronic Business*, 8(4), 54-62.

<https://doi.org/10.5815/ijieeb.2016.04.07>

Dirsehan, T. (2016). Text Mining in the Hospitality Sector To Extend the Motivation Theory, (March). Retrieved from https://www.researchgate.net/profile/Taskin_Dirsehan/publication/298070052_Text_Mining_in_the_Hospitality_Sector_to_Extend_the_Motivation_Theory/links/56e5dde408ae68afa112b709.pdf

Elango, V., & Narayan, G. (2011). Sentiment Analysis for Hotel Reviews. *Proceedings of the Computational Linguistics-Applications Conference*, 231527(March), 45-52. <https://doi.org/10.1051/mateconf/20167503002>

Liu, B. (2012). Sentiment Analysis and Opinion Mining, (May).

Muthia, D. A. (2015). SNIPTEK 2015 OPINION MINING PADA REVIEW HOTEL MENGGUNAKAN ALGORITMA Seminar Nasional Ilmu Pengetahuan dan Teknologi Komputer Nusa Mandiri Seminar Nasional Ilmu Pengetahuan dan Teknologi Komputer Nusa Mandiri, 735-740.

Sipayung, E. M., Maharani, H., & Zefanya, I. (2016). Perancangan Sistem Analisis Sentimen Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier. *Jurnal Sistem Informasi (JSI)*, 8(1), 958-965.

Sodanil, M. (2016). Multi-Language Sentiment Analysis for Hotel Reviews. *MATEC Web of Conferences*, 75, 3002. <https://doi.org/10.1051/mateconf/20167503002>

Tian, X., Tao, R., He, W., & Akula, V. (2016). Mining online hotel reviews: A case study from hotels in China. *AMCIS 2016: Surfing the IT Innovation Wave - 22nd Americas Conference on Information Systems*, (2011), 1-8.

Yordanova, S., & Kabakchieva, D. (2017). Sentiment Classification of Hotel Reviews in Social Media with Decision Tree Learning. *International Journal of Computer Applications*, 158(August), 7. <https://doi.org/10.5120/ijca2017912806>