

Analisis Sentimen Review Produk Kecantikan menggunakan Metode Naïve Bayes

Binti Najibah Agus Ratri¹, Yuita Arum Sari², Indriati³

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya
Email: ¹bintinajibah@student.ub.ac.id, ²yuita@ub.ac.id, ³indriati.tif@ub.ac.id

Abstrak

Masyarakat sedang meminati jual beli dengan menggunakan e-commerce. Di tempat jual beli banyak produk kebutuhan, salah satunya yaitu produk kecantikan. Banyak sekali produk kecantikan yang menawarkan berbagai keunggulan, dengan banyaknya produk tidak lepas dari *review* yang menilai suatu produk tertentu terkait kualitas, keunggulan, kekurangan, dan lainnya. *Review* dari salah satu customer yang pernah menggunakan produk tersebut dapat dijadikan sebagai rekomendasi untuk memilih suatu produk terbaik dan sebagai penentu suatu kualitas produk. Penentuan kualitas produk dapat dilihat dari berbagai komentar atau *review* dari customer untuk melihat apakah produk tersebut menjadi produk best seller atau produk yang kurang diminati. Oleh karena itu dilakukan klasifikasi data *review* produk kecantikan dengan melabeli *review* positif, negatif, dan netral. Pada pengujian ini menggunakan metode Naïve Bayes dengan pembobotan Log TF dan TF, pembobotan TF memiliki nilai akurasi yang lebih baik dibandingkan dengan pembobotan Log TF. Pembobotan TF memiliki akurasi sebesar 55% sedangkan pembobotan Log TF hanya memiliki akurasi sebesar 52%. Sedangkan untuk melabeli *review* menggunakan Kappa Measure, pada pengujian ini Kappa Measure pada setiap penilai memiliki nilai sama yaitu 0.8.

Kata kunci: *Review, Naïve Bayes, Log TF, TF, Kappa Measure.*

Abstract

People are interested in buying and selling using e-commerce. In the place of buying and selling many products, one of which is beauty products. Many beauty products offer various advantages, with many products not being separated from reviews that assess a particular product regarding quality, advantages, disadvantages, and others. A review from a customer who has used the product can be used as a recommendation to choose the best product and as a determinant of product quality. Determination of product quality can be seen from various comments or reviews from customers to see whether the product is a best seller product or a less desirable product. Therefore, beauty product review data classification is carried out by labelling positive, negative, and neutral reviews. In this test using the Naïve Bayes method with TF and TF Log weighting, TF weighting has a better accuracy value than TF Log weighting. TF weighting has an accuracy of 55% while TF Log weighting only has an accuracy of 52%. Meanwhile, to label the review using a kappa measure, in this test the kappa measure on each rater has the same value, namely 0.8.

Keywords: *Review, Naïve Bayes, Log TF, TF, Kappa Measure*

1. PENDAHULUAN

Ecommerce merupakan tempat jual beli segala bentuk produk kebutuhan, salah satunya yaitu produk kecantikan. Banyak sekali produk kecantikan yang menawarkan keunggulannya masing-masing sehingga dapat dijadikan sebagai rekomendasi untuk memilih suatu produk terbaik dan sebagai penentu suatu kualitas

produk. Penilaian suatu produk tidak lepas dari *review* yang dilakukan konsumen dalam menilai terkait kualitas, keunggulan, kekurangan, dan lainnya. Penentuan kualitas produk dapat dilihat dari berbagai komentar atau *review* dari customer untuk melihat apakah produk tersebut menjadi produk best seller atau produk yang kurang diminati.

Penelitian sebelumnya menggunakan Naïve

Bayes, yaitu penelitian dari Wongkar and Angdressey (2019) dalam penelitiannya mereka menggunakan Naive Bayes untuk memprediksi kelas, setelah itu dibandingkan dengan metode lain yaitu SVM dan KNN. Review dikelompokkan menjadi kelas positif, negatif, dan netral. Pada percobaan Naive Bayes menghasilkan tingkat akurasi 80,90%, sedangkan metode lain seperti KNN menghasilkan akurasi 75,58%, dan SVM menghasilkan akurasi 63,99%.

Pada penelitian selanjutnya dilakukan oleh Pugsee, Sombatsri and Juntiwakul (2017) mereka mengimplementasikan metode menggunakan Naive Bayes Classifier sebagai pembelajaran mesin untuk secara otomatis mengklasifikasikan komentar positif atau negatif pada kosmetik produk. Hasil pengujian menunjukkan bahwa klasifikasi aplikasi memiliki akurasi dan presisi tinggi pada positif dan komentar negatif.

Penelitian oleh Gunawan, Pratiwi and Pratama, (2018) melakukan penelitian dengan metode Naive Bayes untuk analisis sentimen pada ulasan produk online, dengan data latih 90% dan data uji 10% menghasilkan akurasi sebesar 77,78%. Data latih mempengaruhi dalam suatu prediksi, karena kualitas yang data tinggi akan menghasilkan akurasi tinggi.

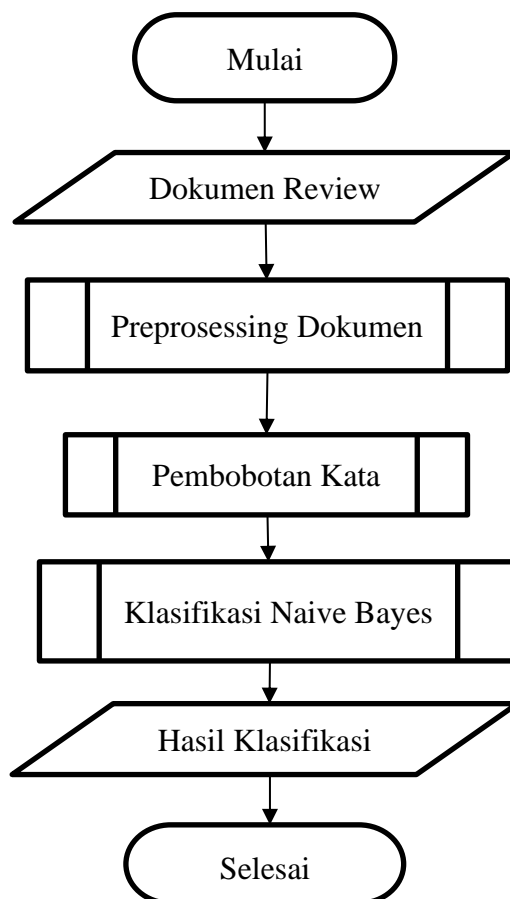
Dari beberapa permasalahan yang dilakukan peneliti sebelumnya, sehingga dalam penelitian ini akan membuat suatu sistem yang menganalisis review produk kecantikan dengan menggunakan metode Naive Bayes. Dengan pemilihan metode tersebut dapat menghasilkan yang terbaik dan dapat dijadikan pertimbangan customer dalam menentukan produk kecantikan.

2. METODE PENELITIAN

2.1 Deskripsi Umum Sistem

Alur penyelesaian sistem dapat dilihat pada Gambar 1. Langkah pertama, yaitu memasukkan dokumen yang akan direview. Untuk data yang digunakan berasal dari situs web www.FemaleDaily.com. Data yang diambil terdiri atas 3 kategori produk, yaitu pelembab wajah, facial wash, dan toner. Dari beberapa kategori diambil 5 produk dari brand yang berbeda. Data yang akan digunakan berupa tabel dengan 7 fitur dan 1.068 data ulasan. Langkah kedua yaitu, dilakukan *pre-processing*, pembobotan, dan klasifikasi menggunakan metode Naive

Bayes. Jumlah kelas untuk penentuan klasifikasi pada sistem dibedakan menjadi 3 kelas, yaitu kelas positif, negatif, dan netral. Sistem ini terdiri atas 2 proses, yaitu proses training dan testing. Hasil akhir penilaian produk nantinya berdasarkan perhitungan nilai probabilitas positif, netral, dan negatif.



Gambar 1. Diagram alir perancangan sistem

2.2 Pre-processing

Tahap *pre-processing* ini merupakan tahapan pertama yang bertujuan untuk mengolah teks yang tidak terstruktur menjadi terstruktur. Dalam *pre-processing* terdiri dari beberapa proses lagi, yaitu *case folding*, *tokenizing*, *filtering*, dan *stemming*. Namun dilakukan perubahan kata alay menjadi kata baku terlebih dahulu sebelum masuk ke tahap *filtering*.

2.3 Term Frequency (TF)

TF merupakan perhitungan yang paling sederhana dalam proses pembobotan suatu kata. Dalam proses ini dengan menghitung jumlah kemunculan *term* dalam setiap

dokumen. Rumus TF dapat dilihat dalam persamaan 1.

$$WTF(t_i, d_j) = f(t_i, d_j) \quad (1)$$

Keterangan :

f : Jumlah kemunculan term

t_i : Term

d_j : Dokumen

2.4 Log frequency (Log TF)

Setiap pembobotan memiliki kekurangan dan kelebihan, begitu pula dengan *Log Frequency*. *Log frequency* dianggap lebih baik dan stabil dibandingkan dengan pembobotan lain karena logaritma digunakan untuk mengatur frekuensi setiap kata dalam dokumen. Rumus *Log Frequency* bisa dilihat pada persamaan 2 (Hendra Suputra, Prebiana and Gorianto, 2021).

$$\text{Log TF} : 1 + \log(TF) \quad (2)$$

Keterangan :

$\text{Log}(tf)$ = Logaritmas basis 10 dari *term frequency*

2.5 Naïve Bayes

Naïve Bayes merupakan metode klasifikasi berdasarkan *Teorema Bayes*, dalam *Teorema Bayes*, jika terdapat suatu permasalahan (Misalnya c dan w) probabilitas dapat diuraikan pada persamaan 3 sebagai berikut.

$$P(c|w) = \frac{P(w*c)*P(c)}{P(w)} \quad (3)$$

Keterangan:

$P(c|w)$: *Posterior*, peluang kemunculan c terhadap w

$P(w|c)$: *Likelihood*, peluang kata w terhadap kelas c

$P(c)$: *Prior*, peluang kemunculan kelas c

$P(w)$: *Evidance*, peluang kemunculan kata w

Pada proses perhitungan klasifikasi, *Evidance* atau peluang kemunculan kata w dapat dihilangkan karena tidak terlalu berpengaruh terhadap hasil klasifikasi dari setiap kelas yang ada. Sehingga

disederhanakan seperti pada persamaan 4 (Rahman and Doewes, 2017).

$$P(c|w_i) = P(c) \times P(w_1|c) \times P(w_2|c) \times P(w_3|c) \times \dots \times P(w_n|c) \quad (4)$$

Pada Persamaan 3 terdapat *prior*, *likelihood*, dan *posterior*. Perhitungan prior dapat dilihat pada Persamaan 5.

$$P(C_j) = \frac{Nc}{N} \quad (5)$$

Keterangan :

Nc : jumlah dokumen pada data latih yang berkelas c

N : jumlah dokumen pada data latih

Untuk perhitungan *likelihood* yaitu, menggunakan model multinomial dengan cara menghitung jumlah kemunculan kata pada setiap dokumen. Dapat dilihat pada persamaan 6.

$$P(w|c) = \frac{\text{count}(w,c)}{\text{count}(c)} \quad (6)$$

Keterangan :

$\text{count}(w|c)$: jumlah kemunculan w pada kelas c

$\text{count}(c)$: jumlah kemunculan semua kata pada kata c

Pada perhitungan model multinomial terdapat permasalahan, yaitu saat kemunculan kata tidak ada sehingga mengakibatkan perhitungan bernilai nol (Kikuchi *et al.*, 2015). Untuk mengatasi permasalahan tersebut digunakan metode *laplace smoothing* dengan cara ditambahkan angka 1 pada kata yang tidak pernah muncul. Perhitungan *laplace smoothing* dapat dilihat pada persamaan 7.

$$P(w, c) = \frac{\text{count}(w,c)+1}{\text{count}(c)+|v|} \quad (7)$$

Keterangan :

$|v|$: jumlah kata unik

3. PENGUJIAN

3.1 Pengujian dengan pembobotan Log TF

Tabel 1. Pengujian dengan pembobotan Log TF

		Actual		
		Negatif	Netral	Positif
	Hasil Prediksi	Negatif	Netral	Positif
		39	6	5
		36	54	28
		18	47	87
Accuracy		56.25%	56.25%	56.25%
Precision		0.78	0.46	0.57
Recall		0.41	0.50	0.73
F-Measure		0.55	0.48	0.64

Berdasarkan Tabel 1. Pengujian menggunakan metode *Naïve Bayes* dengan pembobotan *Log TF* pada pengujian prediksi ke *actual*, untuk prediksi negatif ke *actual* negatif mempunyai nilai sebesar 39, prediksi negatif ke *actual* netral terdapat hasil senilai 6, dan untuk prediksi negatif ke *actual* positif terdapat hasil sebesar 5. Selanjutnya untuk pengujian prediksi netral ke *actual* negatif menghasilkan nilai 36, untuk prediksi netral ke *actual* netral menghasilkan nilai 54, untuk prediksi netral ke *actual* positif menghasilkan nilai 28, untuk prediksi positif ke *actual* negatif menghasilkan nilai 18, untuk prediksi positif ke *actual* netral menghasilkan nilai 47, untuk prediksi positif ke *actual* positif menghasilkan nilai 87. *Accuracy* pada setiap *review* mempunyai hasil yang sama yaitu sebesar 56,25%, sedangkan *precision* pada *review* positif mempunyai hasil 0.57, *review* negatif 0.78, dan *review* netral 0.46. Selanjutnya untuk perhitungan *recall* mempunyai hasil untuk *review* positif 0.73, *review* negatif 0.41, dan *review* netral 0.50. Berikutnya untuk perhitungan *F-Measure* mempunyai hasil *review* positif sebesar 0.64, *review* negatif 0.55, dan *review* netral 0.48.

3.2. Pengujian dengan pembobotan TF

Tabel 2. Pengujian dengan pembobotan TF

		Hasil		
		Negatif	Netral	Positif
	Hasil Prediksi	Negatif	Netral	Positif
		41	7	5
		34	57	28
		18	43	87
Accuracy		57.81%	57.81%	57.81%
Precision		0.77	0.48	0.59
Recall		0.44	0.53	0.73
F-Measure		0.56	0.50	0.65

Berdasarkan Tabel 2. Pengujian menggunakan

metode *Naïve Bayes* dengan pembobotan TF pada pengujian prediksi negatif ke *actual* negatif mempunyai nilai sebesar 41, prediksi negatif ke *actual* netral menghasilkan nilai 7, dan untuk prediksi negatif ke *actual* positif terdapat hasil 5. Selanjutnya untuk pengujian prediksi netral ke *actual* negatif menghasilkan nilai 34, untuk prediksi netral ke *actual* netral menghasilkan nilai 57, untuk prediksi netral ke *actual* positif menghasilkan nilai 28. Kemudian untuk prediksi positif ke *actual* negatif menghasilkan nilai sebesar 18, untuk prediksi positif ke *actual* netral menghasilkan nilai 43, untuk prediksi positif ke *actual* positif menghasilkan nilai 87. *Accuracy* pada setiap *review* mempunyai nilai yang sama yaitu 57,81%, sedangkan *precision* pada positif mempunyai nilai 0.59, *precision* negatif mempunyai nilai 0.77, *precision* netral mempunyai nilai 0.48. Selanjutnya untuk perhitungan *recall* positif menghasilkan nilai 0.73, *recall* negatif menghasilkan nilai 0.44, *recall* netral menghasilkan nilai 0.53. Berikutnya untuk *F-Measure* mempunyai hasil positif sebesar 0.65, hasil *F-Measure* negatif sebesar 0.56, dan hasil *F-Measure* netral menghasilkan nilai sebesar 0.50.

3.3 Evaluasi Relevansi

Evaluasi relevansi pada penelitian ini bertujuan untuk mengukur kesepakatan antara penilai dengan melakukan perhitungan *Kappa Measure*. Perhitungan dilakukan dengan perbandingan antara tiga penilai, yaitu penilai 1 dengan penilai 2, penilai 1 dengan penilai 3, dan penilai 2 dengan penilai 3. Hasil perbandingan penilai 1 dan penilai 2 dapat dilihat pada Tabel 3.

Tabel 3. Perbandingan penilai 1 dan penilai 2

		Penilai 2		
		Positif	Negatif	Netral
	Penilai 1	Positif	Negatif	Netral
		396	7	25
		0	241	68
		28	18	285
		424	266	378
		1068		
<i>Kappa Measure</i>		0,86329588		

Berdasarkan Tabel 3. hasil perbandingan penilai 1 dengan penilai 2 menghasilkan nilai *Kappa Measure* sebesar 0,86329588.

Untuk hasil perbandingan penilai 1 dengan penilai 3 dapat dilihat pada Tabel 4.

Tabel 4. Perbandingan penilai 1 dan penilai 3

		Penilai 3			
		Positif	Negatif	Netral	Total
Penilai 1	Positif	392	6	30	428
	Negatif	1	249	59	309
	Netral	27	19	285	331
	Total				106
		420	274	374	8
Kappa Measure		0,867041199			

Berdasarkan Tabel 4. hasil perbandingan antara penilai 1 dengan penilai 3 menghasilkan nilai *Kappa Measure* 0,867041199.

Kemudian hasil perbandingan antara penilai 2 dan penilai 3 dapat dilihat pada Tabel 5.

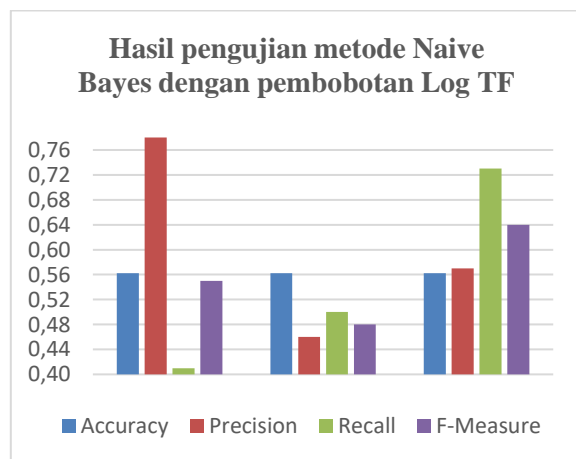
Tabel 5. Perbandingan penilai 2 dan penilai 3

		Penilai 3			
		Positif	Negatif	Netral	Total
Penilai 2	Positif	396	0	28	424
	Negatif	2	227	37	266
	Netral	22	47	309	378
	Total				106
		420	274	374	8
Kappa Measure		0,872659176			

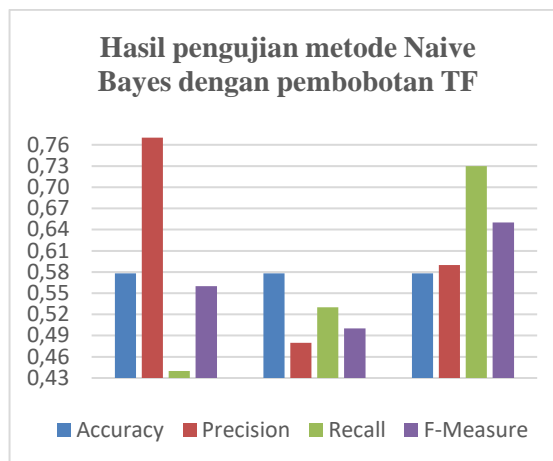
Berdasarkan Tabel 5. hasil perbandingan antara penilai 2 dan penilai 3 menghasilkan nilai *Kappa Measure* sebesar 0,872659176

4. ANALISIS HASIL PENGUJIAN

Setelah dilakukannya pengujian didapatkan hasil seperti pada Gambar 2. dan Gambar 3.



Gambar 2. Hasil pengujian pembobotan Log TF



Gambar 3. Hasil pengujian pembobotan Log TF

Berdasarkan Gambar 2. Hasil pengujian pembobotan Log TF dan Gambar 3. dapat dilihat bahwa hasil pengujian menggunakan metode *Naive Bayes* dengan pembobotan TF menghasilkan nilai akurasi lebih baik dibandingkan dengan hasil pengujian menggunakan metode *Naive Bayes* dengan pembobotan *Log TF*. Nilai akurasi dengan menggunakan pembobotan TF memiliki nilai sebesar 57,81%, sedangkan nilai akurasi dengan menggunakan pembobotan *Log TF* memiliki nilai sebesar 56,25%. Hal ini terjadi karena perbedaan hasil *recall* pada kelas negatif, penggunaan pembobotan TF yang lebih sederhana ternyata mampu memberikan hasil sensitifitas yang lebih tinggi daripada penggunaan Log TF pada kelas negatif, sehingga berpengaruh pada hasil f-measure dan akurasi pada seluruh hasil pengujian.

Tabel 6. Hasil klasifikasi dengan hasil sebenarnya

Review	Hasil Klasifikasi	Hasil Sebenarnya
Moistnya ringan dan sedikit nimbulin whitecast saat pertama di apply. Tapi setelah beberapa menit lgsg menyerap dan keliatan natural bgt ❤️ First pake, jd belum nimbulin efek apapun selain ngemoist kulit dgn baik ❤️ Repurchase? i think yay 🥰	Positif	Netral
Suka bgt! Pelembap nya enak berasa lgsg lembap. Ga berat cocok buat krim pagi. Untuk cerahin atau apapun itu aku ga rasain sih. Cocok aja di aku buat	Positif	Netral

lembapin muka lgsg enak rasanya. Apalagi ada spf juga lumayan ga perlu pake sunscreen klo dirumah doang

	Positif	Netral
3 hari pertama pemakaian wajah aku jadi berjerawat. tapi lama kelamaan udh ga lg berjerawat. kalaupun berjerawat, jerawatnya agak lebih calm. trus kalo make nya jgn kebanyakan cukup sebesar bini jagung. karna kalau kebanyakan, bkin wajah kilang minyak 😊. but, udah 1 bulan lebih aku make. wajah aku udah lumayan putih dan bersih. biasanya aku makenya stelah make micellar water, sama toner. *update pemakaian 2 bulan w ngerasa ada efek brighteningnya trus jerawat w udah mulai memudar dan bersih recomended banget sih. ini aja ga abis2 🙏🙏🙏		

Berdasarkan Tabel 6. dapat diketahui terjadi beberapa kesalahan pada klasifikasi yang menyebabkan turunnya hasil akurasi. Contoh pada review pertama seharusnya diklasifikasi sebagai kelas netral, namun hasilnya positif. Pada review pertama kalimat “sedikit nimbunin whitecast saat pertama di apply. Tapi setelah beberapa menit lgsg menyerap dan keliatan natural bgt” mengindikasikan termasuk pada kelas netral. hal ini terjadi karena jumlah data yang tidak seimbang pada pengujian, kelas positif yang memiliki jumlah data yang lebih banyak, memiliki peluang lebih besar untuk diklasifikasikan, sehingga review yang seharusnya netral diklasifikasi sebagai positif, selain itu kesamaan review antara positif dan netral pada beberapa data, contoh pada tabel 6.6, menjadi faktor yang menyebabkan turunnya performa sistem.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil pengujian dan analisis yang telah dilakukan dalam penelitian ini yang menggunakan metode *Naïve Bayes* dengan pembobotan *Log TF* dan *TF* menghasilkan kesimpulan sebagai berikut : Metode *Naïve Bayes* mampu digunakan untuk analisis sentimen review produk kecantikan. Dalam penelitian ini melakukan

klasifikasi dengan menggunakan metode *Naïve Bayes* dan memberi label positif, negatif, netral pada review, dengan pembobotan menggunakan *Log TF* dan *TF*. Dari hasil akurasi dapat diketahui bahwa pembobotan *TF* lebih baik dibandingkan dengan pembobotan *Log TF*. Pembobotan *TF* menghasilkan akurasi sebesar 57,81%, sedangkan pembobotan dengan *Log TF* hanya menghasilkan akurasi sebesar 56,25%.

Saran dari kekurangan yang telah dilakukan dalam penelitian ini yaitu, Metode *Naïve Bayes* dapat digunakan sebagai analisis sentimen produk kecantikan namun hasil akurasi masih terbilang rendah, maka dari itu perlu adanya penelitian dengan metode lain atau penggunaan pembobotan lainnya sehingga dapat dijadikan pembandingan dan atau meningkatkan kinerja sistem.

6. DAFTAR PUSTAKA

- Gunawan, B., Pratiwi, H. S. and Pratama, E. E. (2018) ‘Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes’, *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, 4(2), p. 113. doi: 10.26418/jp.v4i2.27526.
- Hendra Suputra, I. P. G., Prebiana, K. D. and Gorianto, F. O. (2021) ‘Perbandingan Jenis TF terhadap Hasil Evaluasi Information Retrieval’, *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, 8(2), p. 207. doi: 10.24843/jlk.2019.v08.i02.p13.
- Kikuchi, M. et al. (2015) ‘Confidence interval of probability estimator of Laplace smoothing’, *ICAICTA 2015 - 2015 International Conference on Advanced Informatics: Concepts, Theory and Applications*. doi: 10.1109/ICAICTA.2015.7335387.
- Pugsee, P., Sombatsri, P. and Juntiwakul, R. (2017) ‘Satisfactory analysis for cosmetic product review comments’, in *ACM International Conference Proceeding Series*. doi: 10.1145/3089871.3089890.
- Wongkar, M. and Angdresey, A. (2019) ‘Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler: Twitter’, in *Proceedings of 2019 4th International Conference on Informatics and Computing, ICIC 2019*. doi:

10.1109/ICIC47613.2019.8985884.