

Analisis Sentimen Opini Film Menggunakan Metode *Naïve Bayes* dan *Lexicon Based Features*

Arifin Kurniawan¹, Indriati², Sigit Adinugroho³

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya
Email: ¹arifinkurns@gmail.com, ²indriati.tif@ub.ac.id, ³sigit.adinu@ub.ac.id

Abstrak

Perkembangan teknologi informasi yang semakin pesat mengakibatkan banyak orang yang menulis opini mereka di media sosial seperti pada forum KASKUS. KASKUS merupakan situs forum online yang menyediakan tempat untuk mencari informasi dan berbagi hobi. Salah satunya adalah forum Movies yang berisi diskusi mengenai suatu film yang telah ditonton. Pengguna menuliskan opininya mengenai suatu film apakah film tersebut bagus atau jelek. Opini-opini tersebut dapat dianalisis untuk mengetahui bagaimana tanggapan pengguna tentang film tersebut agar menghasilkan output yang bermanfaat bagi pembuat film dengan melakukan analisis sentimen untuk mengklasifikasikan opini ke dalam kelas positif atau kelas negatif. Analisis sentimen dilakukan menggunakan metode *Naïve Bayes* untuk klasifikasi dan *Lexicon Based Features* untuk pembobotan nilai sentimen suatu kata. Proses yang dilakukan dimulai dari *text preprocessing*, *term weighting*, *Naïve Bayes training*, dan *Naïve Bayes testing* dengan pembobotan *lexicon based features* menggunakan kamus Barasa. Berdasarkan hasil pengujian yang dilakukan, dengan menggunakan metode *Naïve Bayes* dan *Lexicon Based Features* didapatkan nilai *accuracy*, *precision*, *recall*, dan *f-measure* sebesar 0,8, 0,8, 0,8 dan 0,8. Sedangkan dengan menggunakan metode *Naïve Bayes* tanpa pembobotan *Lexicon Based Features* didapatkan nilai *accuracy*, *precision*, *recall*, dan *f-measure* sebesar 0,95, 1, 0,9 dan 0,9474. Sehingga penggunaan metode *Naïve Bayes* dengan *Lexicon Based Features* masih belum dapat memberikan hasil lebih baik.

Kata kunci: KASKUS, film, analisis sentimen, *Naïve Bayes*, *Lexicon Based Features*, Barasa

Abstract

The rapid development of information technology has resulted in many people writing their opinions on social media as in the KASKUS forum. KASKUS is an online forum site that provides a place to find information and share hobbies. One is called Movies forum which contains discussions about a movie that has been watched. Users writes their opinion about a film whether the film is good or bad. These opinions can be analyzed to determine how the user feedback about the film in order to produce useful output for the filmmaker by perform sentiment analysis to classify opinions into positive or negative classes. The process of sentiment analysis was performed using methods Naïve Bayes for classification and Lexicon Based Features to weight the sentiment value of a word. The process starts from text preprocessing, term weighting, Naïve Bayes training, and Naïve Bayes testing with Lexicon Based Features weighting using Barasa's lexicon. Based on the results of tests performed, using Naïve Bayes and Lexicon Features Based method, the values of accuracy, precision, recall, and f-measure were 0.8, 0.8, 0.8 and 0.8. While using the Naïve Bayes method without Lexicon Based Features, the values of accuracy, precision, recall, and f-measure are 0.95, 1, 0.9 and 0.9474. So, the use of Naïve Bayes and Lexicon Based Features methods still cannot provide better results.

Keywords: KASKUS, movie, sentiment analysis, *Naïve Bayes*, *Lexicon Based Features*, Barasa

1. PENDAHULUAN

Film merupakan media komunikasi yang bersifat audio visual (suara dan gambar) yang mana berisi pesan tersirat yang disampaikan oleh pembuat film kepada sekelompok orang yang menyaksikannya (Effendy, 1986). Pesan yang tersirat pada suatu film berpengaruh kepada pemikiran penonton terhadap maksud dari film tersebut. Dengan menyaksikannya, penonton seakan-akan dibawa masuk mengikuti jalan cerita sehingga dapat mempengaruhi persepsi penonton untuk menyampaikan opini mereka tentang bagaimana alur, latar, penokohan, dan *ending* cerita.

Opini atau pendapat adalah pandangan seseorang terhadap suatu persoalan (Abdullah, 2001). Opini yang disampaikan oleh seseorang terhadap suatu hal yang sama dapat menyebabkan munculnya penilaian yang berbeda. Semakin pesatnya perkembangan teknologi informasi, semakin banyak orang yang menulis opini mereka tentang sebuah produk atau layanan di media sosial (Ratnawati & Winarko, 2018). Opini-opini tersebut dituliskan pada sebuah *blog* atau *website*. Beberapa *blog* dan *website* yang mengulas tentang suatu film menggunakan Bahasa Indonesia antara lain Cinemags Magazine, Cenayang Film, Watchmen ID, dan Forum Movies Kaskus. *Blog* dan *website* tersebut mengulas tentang film yang berupa opini dari sudut pandang penulis. Opini-opini tersebut dapat dianalisis untuk mengetahui bagaimana tanggapan penulis tentang film tersebut yang menghasilkan *output* yang bermanfaat bagi pembuat film yaitu dengan melakukan analisis sentimen.

Penelitian terdahulu perihal analisis sentimen menggunakan metode *naïve bayes* terhadap data opini di *twitter* di beberapa situs *online marketplace* di Indonesia. Dari pengujian yang dilakukan, didapatkan rata-rata akurasi sebesar 93.33% (Artanti, Syukur, Prihandono, & Setiadi, 2018). Penelitian berikutnya yaitu dengan melakukan implementasi algoritme *naïve bayes* terhadap data film pada media sosial *Twitter*. Dari hasil pengujian, didapatkan nilai akurasi sebesar 90% (Ratnawati, 2018).

Pada penelitian yang dilakukan oleh Siddiqua, et al., (2017) yang melakukan analisis sentimen dengan menggabungkan *rule-based classifier* dengan *ensemble features* dan metode *machine learning*. Dari pengujian yang dilakukan, penggunaan *lexicon based features* dapat berperan penting pada analisis sentimen. Penelitian selanjutnya yang dilakukan oleh Suryadi & Winarko (2017) yang menggunakan metode *naïve bayes classifier* dan pendekatan *lexicon based* untuk analisis sentimen *review* hotel. Dari pengujian yang dilakukan, didapatkan rata-rata akurasi sebesar 89,95%, rata-rata presisi sebesar 98,55%, dan rata-rata *f-measure* sebesar 91,07%.

Penelitian berikutnya perihal analisis sentimen yang dilakukan oleh Lestari, et al. (2019) yang menggunakan metode *Learning Vector Quantization* (LVQ) dan *lexicon based features* untuk klasifikasi video *clickbait* pada YouTube. Berdasarkan evaluasi yang dilakukan, penggunaan metode *Learning Vector Quantization* didapatkan akurasi sebesar 54,54%, *precision* sebesar 1, *recall* sebesar 0,1667, dan *f-measure* sebesar 0,2858. Sedangkan menggunakan metode *Learning Vector Quantization* dan *lexicon based features* didapatkan akurasi sebesar 90,91%, *precision* sebesar 0,8571, *recall* sebesar 1, dan *f-measure* sebesar 0,9231. Hal tersebut membuktikan penggunaan *lexicon based features* mampu meningkatkan akurasi sistem.

Berdasarkan uraian di atas, penggunaan metode *naïve bayes* dan *lexicon based features* diharapkan mampu memberikan hasil klasifikasi yang baik dan akurat.

2. DASAR TEORI

2.1. KASKUS

KASKUS merupakan situs forum online tempat Berbagi Hobi & Berkomunitas pertama serta terbesar di Indonesia. KASKUS menyediakan tempat untuk mencari informasi, berbagi hobi, bergabung dan berinteraksi dengan komunitas, bertransaksi jual beli pada empat platform yaitu: Forum, Jual Beli (JB), TV, dan Podcast. Pada Forum terdapat beberapa kategori dan subforum di setiap kategori. Salah satunya adalah Movies. Pada setiap forum dan subforum terdapat tulisan atau buah pikiran seseorang yang disebut *thread*. Di dalam

sebuah *thread*, pengguna yang terdaftar atau yang disebut Kaskuser dapat melakukan komentar atau balasan.

2.2. Text Mining

Text mining merupakan prosedur menemukan pola atau informasi yang pada awalnya tidak tampak dalam sebuah dokumen teks atau sumber tertentu yang nantinya akan digunakan sebagai informasi yang bermanfaat untuk tujuan tertentu (Adeva & Calvo, 2006). Penggunaan *text mining* dapat menyelesaikan masalah seperti analisis sentimen, *document clustering*, *document classification*, *information extraction*, *information retrieval*, dan *web mining*.

2.3. Analisis Sentimen

Analisis sentimen merupakan bidang ilmu yang menganalisis opini, sentimen, evaluasi, penilaian, sikap, dan emosi terhadap suatu entitas seperti produk, layanan, organisasi, individu, isu, peristiwa, dan topik (Liu, 2012). Analisis sentimen berfokus pada opini yang mengekspresikan sentimen positif atau negatif.

2.4. Text Preprocessing

Text preprocessing adalah tahap permulaan pada proses *text mining* untuk memperoleh data yang siap untuk diproses (Indriati & Ridok, 2017). Urutan pada *text preprocessing* yaitu: *stemming*, *tokenizing*, dan *filtering*. Tahapan *text preprocessing* menghasilkan kumpulan kata yang nantinya dijadikan sebagai indeks.

2.4.1. Stemming

Stemming ialah proses mengganti bentuk kata menjadi kata dasarnya. Setiap kata berimbuhan (awalan dan akhiran) akan dihilangkan dan membentuk kata dasar sehingga proses *text mining* dapat lebih optimal. Pada penelitian ini digunakan *library* Python Sastrawi *Stemmer*.

2.4.2. Tokenizing

Tokenizing ialah proses memecah kata-kata dalam suatu dokumen. Dalam proses ini dilakukan penghapusan terhadap karakter *whitespace* karena karakter-karakter tersebut tidak berpengaruh terhadap *text preprocessing*.

2.4.3. Filtering

Filtering merupakan tahapan penyaringan kata dari hasil proses *tokenizing*. Penelitian ini menggunakan metode *stoplist*. *Stoplist* adalah kumpulan kata yang tidak penting yang dapat dihilangkan dengan pendekatan *bag-of-words*. *Stopword list* yang digunakan yaitu *stopword list* Tala dengan penambahan *keyword emoticon* KASKUS.

2.5. Term Weighting

Term weighting ialah proses mengubah indeks atau fitur dari hasil *preprocessing* yang berupa data kata-kata menjadi data angka dengan memberi nilai atau bobot untuk setiap kata. Hasil dari proses *term weighting* dapat digunakan untuk proses klasifikasi. Metode yang digunakan yaitu *raw-term frequency weighting*. *Raw-term frequency weighting* adalah bobot suatu kata dalam suatu dokumen berdasarkan jumlah kemunculan kata tersebut dalam dokumen. Dapat dilihat pada Persamaan (1).

$$w_{t,d} = tf_{t,d} \quad (1)$$

dimana $w_{t,d}$ adalah bobot kemunculan kata t pada dokumen d dan $tf_{t,d}$ adalah kemunculan kata t pada dokumen d .

2.6. Klasifikasi

Klasifikasi merupakan proses memasukan suatu data yang kelasnya belum diketahui berdasarkan data yang kelasnya sudah didefinisikan sebelumnya. Proses klasifikasi terbagi dalam dua fase, yaitu *learning step* dan *classification step*. Pada *learning step*, dilakukan pembuatan model klasifikasi dengan menggunakan algoritme yang mempelajari data yang kelasnya sudah didefinisikan sebelumnya atau data latih. Karena kelasnya sudah diketahui sebelumnya, tahap ini disebut *supervised learning*. Selanjutnya model hasil *learning step* digunakan pada *classification step* untuk memprediksi suatu data yang kelasnya belum diketahui (Han, et al., 2011).

2.7. Naïve Bayes Classifier

Naïve bayes classifier merupakan metode klasifikasi berbasis statistik berdasarkan teorema Bayes untuk mengklasifikasikan suatu data ke dalam kelas yang sudah ditentukan sebelumnya (Han, et al., 2011). Disebut 'naïve' karena nilai dari suatu atribut tidak berpengaruh

atau independen terhadap nilai dari atribut lainnya yang disebut *conditional independence*. *Naïve bayes classifier* menunjukkan tingginya akurasi dan kecepatan ketika diimplementasikan ke data yang besar dibandingkan dengan performa dari *decision tree* dan beberapa algoritme klasifikasi *neural network* (Han, et al., 2011). Persamaan (2) menunjukkan perhitungan *naïve bayes classifier* secara umum.

$$P(c|w) = \frac{P(w|c)*P(c)}{P(w)} \tag{2}$$

dimana $P(c|w)$ adalah *posterior* atau peluang kelas c terhadap kata w , $P(w|c)$ adalah *likelihood* atau peluang kata w terhadap kelas c , $P(c)$ adalah *prior* atau peluang kemunculan kelas c , dan $P(w)$ adalah *evidence* atau peluang kemunculan kata w .

Pada Persamaan (2) terdapat perhitungan nilai *likelihood*, *prior*, dan *evidence*. Perhitungan nilai *likelihood* menggunakan model multinomial. Model ini memperhitungkan jumlah kemunculan setiap kata pada sebuah dokumen. Dapat dilihat pada Persamaan (3).

$$P(w|c) = \frac{\text{count}(w,c)}{\text{count}(c)} \tag{3}$$

dimana $\text{count}(w,c)$ ialah banyaknya kemunculan kata w pada kelas c dan $\text{count}(c)$ ialah banyaknya kemunculan semua kata pada kelas c .

Permasalahan yang sering ditemukan pada perhitungan model *multinomial* ialah dimana adanya suatu kata yang tidak pernah muncul akan mengakibatkan perhitungan bernilai nol yang disebut *zero frequency problem* (Kikuchi, et al., 2015). Cara menangani masalah tersebut, dilakukan *laplace smoothing*. *Laplace smoothing* mengatasi masalah tersebut dengan melakukan penambahan nilai 1 pada kata tersebut atau pada pembilang sehingga dianggap pernah muncul sekali dan penambahan kata unik pada penyebut. Perhitungan *laplace smoothing* dapat dilihat pada Persamaan (4).

$$P(w|c) = \frac{\text{count}(w,c)+1}{\text{count}(c)+|v|} \tag{4}$$

dimana $|v|$ adalah jumlah kata unik. Untuk menghitung *prior* ditunjukkan pada Persamaan (5).

$$P(c_j) = \frac{N_c}{N} \tag{5}$$

dimana N_c adalah jumlah dokumen pada data latih yang berkelas c dan N adalah jumlah dokumen pada data latih. Untuk menghitung nilai *evidence* dapat dilihat pada Persamaan (6).

$$P(w) = \frac{|w|}{|S|} \tag{6}$$

dimana $|w|$ adalah jumlah kata w yang muncul dan $|S|$ merupakan jumlah keseluruhan kata yang muncul dalam seluruh dokumen.

2.8. Lexicon Based Features

Lexicon based features ialah fitur atau kata yang telah diberi bobot berdasarkan kamus atau *lexicon*. Pemberian bobot dilakukan untuk setiap kata yang termasuk sentimen positif atau termasuk sentimen negatif. Tujuan penggunaan *lexicon based features* yaitu untuk menentukan orientasi sentimen suatu kata. Kamus atau *lexicon* yang digunakan yaitu kamus Barasa milik David Moeljadi. Barasa adalah SentiWordnet berbahasa Indonesia yang berasal dari gabungan Bahasa WordNet v1.1 dan SentiWordNet 3.0 Bahasa Inggris. Struktur dari Barasa yaitu: *synset*, *language*, *goodness*, *lemma*, *PosScore*, dan *NegScore*. Kamus Barasa didapat dari proses *generate file barasa.py* menggunakan *Anaconda Prompt*.

Nilai yang dijadikan bobot adalah nilai sentimen yang dihitung dengan perhitungan jumlah dari nilai *PosScore* atau *NegScore* dibagi dengan nilai total dari *PosScore* dan *NegScore* untuk pembobotan masing-masing kata. Dapat dilihat pada Persamaan (7) dan Persamaan (8).

$$\text{sum_PosScore}(kata) = \sum_{i=1}^n \text{PosScore}(kata) \tag{7}$$

$$\text{sum_NegScore}(kata) = \sum_{i=1}^n \text{NegScore}(kata) \tag{8}$$

Perhitungan nilai total dari *PosScore* dan *NegScore* dari suatu kata dapat dilihat pada Persamaan (9).

$$\text{total_score} = \text{sum_PosScore} + \text{sum_NegScore} \tag{9}$$

Perhitungan nilai sentimen suatu kata pada kelas positif dan kelas negatif dapat dilihat dalam Persamaan (10) dan Persamaan (11).

$$\text{senti_score_pos}(kata) = \frac{\text{sum_PosScore}(kata)}{\text{total_score}(kata)} \tag{10}$$

$$\text{senti_score_neg}(kata) = \frac{\text{sum_NegScore}(kata)}{\text{total_score}(kata)} \tag{11}$$

Nilai sentimen suatu kata berupa bilangan real dengan interval 0 sampai 1. Jika nilai sentimen suatu kata mendekati 1, maka

kata tersebut memiliki sentimen yang semakin positif. Sedangkan jika mendekati 0, maka kata tersebut memiliki sentimen yang semakin negatif. Nilai sentimen tersebut nantinya akan diintegrasikan ke dalam perhitungan *posterior* untuk setiap kelas positif dan kelas negatif. Perhitungan tersebut ditunjukkan pada Persamaan (12) (Goel, et al., 2017).

$$P(c|w) = \frac{P(W|C)+senti_score) \cdot P(c)}{P(w)+senti_score} \quad (12)$$

2.9. Evaluasi

Evaluasi bertujuan untuk menilai hasil uji coba sistem yang dibuat apakah telah sesuai antara hasil sistem analisis sentimen dengan hasil sebenarnya. Metode evaluasi yang digunakan adalah evaluasi temu kembali tak berperingkat menggunakan metode *confusion matrix* dan evaluasi relevansi menggunakan *kappa measure*.

2.9.1. Confusion Matrix

Confusion matrix ialah tabel yang digunakan untuk menganalisis seberapa baik keakuratan sebuah metode klasifikasi untuk memprediksi kelas suatu data (Han, et al., 2011). Pada penelitian ini, data diklasifikasikan menjadi positif dan negatif atau disebut klasifikasi biner. Tabel 1 menunjukkan tabel *confusion matrix*.

Tabel 1. *Confusion Matrix*

Kelas	Diklasifikasikan sebagai kelas positif	Diklasifikasikan sebagai kelas negatif
positif	<i>true positive</i> (tp)	<i>false negative</i> (fn)
negatif	<i>false positive</i> (fp)	<i>true negative</i> (tn)

Berdasarkan Tabel 1, dapat dilakukan perhitungan *accuracy*, *precision*, *recall*, dan *f-measure*.

Accuracy adalah persentase dari data uji yang kelasnya diklasifikasikan secara tepat oleh sistem sesuai dengan kelas aslinya. Perhitungan *accuracy* dapat dilihat pada Persamaan (13).

$$accuracy = \frac{tp+tn}{tp+tn+fp+fn} \quad (13)$$

Precision adalah nilai keserasian antara kelas data yang dihasilkan sistem dengan kelas sebenarnya. Perhitungan *precision* dapat dilihat pada Persamaan (14).

$$precision = \frac{tp}{tp+fp} \quad (14)$$

Recall adalah nilai dari ketepatan banyaknya data yang berhasil dihasilkan sistem berdasarkan kelas sebenarnya. Perhitungan *recall* dapat dilihat pada Persamaan (15).

$$recall = \frac{tp}{tp+fn} \quad (15)$$

F-measure adalah ukuran hubungan timbal balik antara *precision* dan *recall*. *F-measure* juga disebut sebagai alternatif dengan menggabungkan *precision* dan *recall* menjadi satu ukuran evaluasi. Perhitungan *f-measure* dapat dilihat pada Persamaan (16).

$$F = \frac{2 \times precision \times recall}{precision + recall} \quad (16)$$

2.9.2. Evaluasi Relevansi

Evaluasi relevansi dilakukan untuk menilai apakah dokumen-dokumen yang dijadikan *dataset* relevan dengan menghitung nilai *kappa measure*. *Kappa measure* digunakan untuk mengukur kesepakatan antara para penilai (Han, et al., 2011). Perhitungan *kappa measure* dapat dilihat pada Persamaan (17).

$$Kappa = \frac{P(A)-P(E)}{1-P(E)} \quad (16)$$

dimana $P(A)$ adalah proporsi banyaknya penilai yang sama dan $P(E)$ adalah proporsi banyaknya penilai yang sama karena kebetulan.

Kappa bernilai 1 jika semua penilai selalu setuju, bernilai 0 jika penilai hanya setuju dengan nilai yang diberikan berdasarkan peluang, dan bernilai negatif jika penilai lebih buruk dari relevansi secara acak. Nilai *kappa* lebih dari 0,8 termasuk kesepakatan yang baik, nilai *kappa* antara 0,67 sampai 0,8 termasuk kesepakatan yang cukup, dan nilai *kappa* dibawah 0,67 termasuk kesepakatan yang diragukan.

3. METODOLOGI

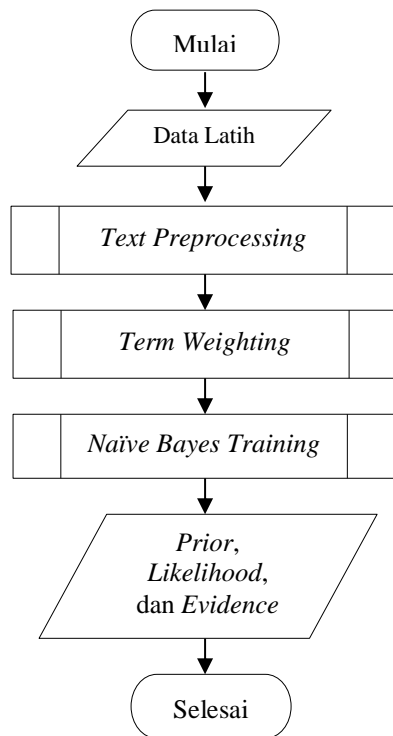
3.1. Pengumpulan Data

Data opini film yang dikumpulkan berupa data primer yang merupakan balasan dari sebuah *thread* atau disebut komentar yang berasal dari Forum Kaskus kategori Movies dengan judul *thread* “[Thread Review Film] The Last Film You Saw - Good/bad/Biasa? MASUK!!” part 3 dan part 2. Komentar yang dipilih untuk dijadikan *dataset* adalah komentar dengan rentang waktu tahun 2011 sampai tahun 2019. Kriteria pemilihan komentar yaitu

komentar yang berbahasa Indonesia, menyertakan judul film, dan bukan balasan dari pengguna yang mengutip komentar yang berisi *review* film. Data yang dikumpulkan sebanyak 150 data. *Dataset* tersebut dibagi menjadi data latih sebanyak 130 dengan pembagian 65 kelas positif dan 65 kelas negatif, dan data uji sebanyak 20 dengan pembagian 10 kelas positif dan 10 kelas negatif. *Dataset* selanjutnya disimpan dalam file *text* (.txt) dengan *tab delimited* sebagai pemisah antara ulasan dengan kelas.

3.2. Diagram Alir Sistem

Proses yang terjadi pada sistem dibagi dalam dua fase, yakni: fase *training* dan fase *testing*. Gambar 1 menunjukkan tahapan yang terjadi pada fase *training*.

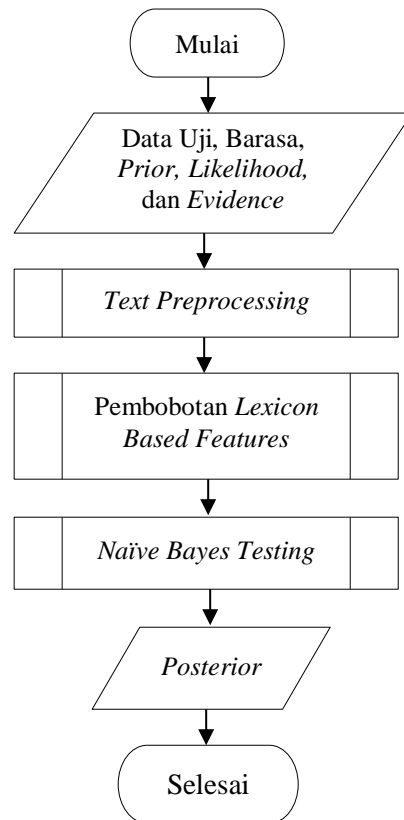


Gambar 1. Fase *training*

Berdasarkan Gambar 1, proses pada fase *training* diawali dengan masukan berupa data latih. Selanjutnya data latih dilakukan proses *text preprocessing*, dimana pada proses tersebut dilakukan *stemming*, *tokenizing*, dan *filtering*. Selanjutnya hasil dari *text processing* dilakukan pembobotan kata atau *term weighting* dengan menggunakan metode pembobotan *raw-term frequency*. Kata yang telah diberi bobot selanjutnya digunakan dalam fase *training* menggunakan metode *Naïve Bayes*, yaitu pada perhitungan nilai *prior*, *likelihood*, dan

evidence. Nilai tersebut akan digunakan sebagai masukan pada fase *testing*.

Tahapan yang terjadi pada fase *testing* dapat dilihat pada Gambar 2.



Gambar 2. Fase *testing*

Berdasarkan Gambar 2, proses pada fase *testing* dimulai dengan masukan data uji, kamus Barasa, dan nilai *prior*, *likelihood*, dan *evidence*. Proses awal yang dilakukan adalah *text preprocessing*, dimana pada proses tersebut dilakukan *stemming*, *tokenizing*, dan *filtering*. Selanjutnya dilakukan pembobotan *lexicon based features* berdasarkan kata yang tersedia di kamus Barasa dengan melakukan perhitungan nilai sentimen. Selanjutnya dilakukan proses klasifikasi *Naïve Bayes*, yaitu dengan menghitung nilai *posterior* sebagai keluaran dari fase *testing*. Nilai *posterior* tertinggi akan menentukan kelas dari data uji.

4. PENGUJIAN DAN ANALISIS

4.1. Evaluasi Temu Kembali Tak Berperingkat

Evaluasi temu kembali tak berperingkat akan dilakukan dengan dua skenario, skenario pertama yaitu klasifikasi menggunakan metode *Naïve Bayes* dengan pembobotan *lexicon based features* dan skenario kedua adalah klasifikasi

menggunakan metode *Naïve Bayes* tanpa pembobotan *lexicon based features*.

Pengujian ini menggunakan 150 data dengan pembagian 130 data latih yang terdiri atas 65 kelas positif dan 65 kelas negatif, dan 20 data uji yang terdiri atas 10 kelas positif dan 10 kelas negatif.

4.1.1. Skenario dengan Pembobotan Lexicon Based Features

Tabel 2 memperlihatkan hasil pengujian dengan pembobotan *lexicon based features*.

Tabel 2. Hasil Pengujian dengan Pembobotan Lexicon Based Features

Accuracy	0,8
Precision	0,8
Recall	0,8
F-Measure	0,8

4.1.2. Skenario tanpa Pembobotan Lexicon Based Features

Tabel 3 memperlihatkan hasil pengujian tanpa pembobotan *lexicon based features*.

Tabel 3. Hasil Pengujian tanpa Pembobotan Lexicon Based Features

Accuracy	0,95
Precision	1
Recall	0,9
F-Measure	0,9474

Berdasarkan Tabel 2 dan Tabel 3, metode *Naïve Bayes* dengan pembobotan *lexicon based features* menghasilkan nilai *accuracy*, nilai *precision*, nilai *recall*, dan nilai *f-measure* yang lebih buruk dibandingkan metode *Naïve Bayes* tanpa pembobotan *lexicon based features*. Hal tersebut disebabkan adanya suatu kata yang memiliki nilai sentimen positif jauh lebih besar dibandingkan dengan nilai sentimen negatif pada dokumen yang kelas sebenarnya negatif, dan sebaliknya. Adanya perbedaan antara nilai sentimen positif dengan nilai sentimen negatif yang sangat besar dapat mengakibatkan kesalahan klasifikasi.

4.2. Evaluasi Relevansi

Evaluasi relevansi dilakukan untuk mengukur kesepakatan antara tiga penilai

dengan melakukan perhitungan *kappa-measure*. Perhitungan dilakukan terhadap 3 perbandingan, yaitu antara penilai 1 dengan penilai 2, penilai 1 dengan penilai 3, dan penilai 2 dengan penilai 3. Tabel 4. memperlihatkan hasil perbandingan antara ketiga penilai

Tabel 4. Hasil Perbandingan Ketiga Penilai

Penilai 1 dengan Penilai 2	1
Penilai 1 dengan Penilai 3	0,9867
Penilai 2 dengan Penilai 3	0,9867

Berdasarkan Tabel 4, perbandingan nilai *kappa-measure* antara 3 penilai masing-masing memiliki nilai *kappa-measure* yang lebih besar dari 0,8. Hal tersebut membuat ketiga perbandingan termasuk kesepakatan yang sangat baik dan dapat dikatakan dokumen-dokumen yang digunakan sebagai *dataset* yang relevan.

5. KESIMPULAN DAN SARAN

Penggunaan metode *Naïve Bayes* dan *lexicon based features* untuk analisis sentimen opini film dapat dilakukan dengan melalui beberapa tahapan dimulai dari *text preprocessing*, menghitung nilai *term weighting*, dan klasifikasi *Naïve Bayes* dengan pembobotan nilai sentimen berdasarkan kamus Barasa.

Pengujian metode *Naïve Bayes* dan *lexicon based features* untuk analisis sentimen opini film menghasilkan nilai *accuracy* sebesar 0,8, *precision* sebesar 0,8, *recall* sebesar 0,8, dan *f-measure* sebesar 0,8. Hal tersebut dikarenakan adanya suatu kata yang memiliki nilai sentimen positif jauh lebih besar dibandingkan dengan nilai sentimen negatif pada dokumen yang kelas sebenarnya negatif, dan sebaliknya. Adanya perbedaan antara nilai sentimen positif dengan nilai sentimen negatif yang sangat besar dapat mengakibatkan kesalahan klasifikasi. Sedangkan penggunaan metode *Naïve Bayes* tanpa *lexicon based features* untuk analisis sentimen opini film menghasilkan nilai *accuracy* sebesar 0,95, *precision* sebesar 1, *recall* sebesar 0,9, dan *f-measure* sebesar 0,9474 yang mana hasil tersebut lebih baik dibandingkan dengan menggunakan *lexicon based features*.

Saran dari kekurangan yang ada untuk penelitian di masa mendatang antara lain:

Penggunaan *postag* untuk mencocokkan bobot kata sesuai dengan *synset* pada kamus Barasa, penambahan bobot positif dan bobot negatif dengan menggunakan *emoticon* Kaskus, perbaikan kata pada data menjadi kata baku agar sesuai dengan kata pada kamus Barasa, dan penggunaan *bigram* untuk mendeteksi kata negasi seperti ‘tidak menarik’.

6. DAFTAR PUSTAKA

- Abdullah, A., 2001. *Press Relations: Kiat Berhubungan Dengan Media Massa*. Bandung: Remaja Rosdakarya.
- Adeva, J. J. G., & Calvo, R. 2006. *Mining Text with Pimiento*. IEEE Internet Computing, 10(4), 27–35.
- Barasa Indonesian SentiWordNet, 2018. [kode program] GitHub. Tersedia di: <<https://github.com/neocl/barasa>>
- Effendy, O. U., 2008. *Dinamika Komunikasi*. Bandung: Remaja Rosdakarya.
- Goel, A., Gautam, J., & Kumar, S., 2017. *Real Time Sentiment Analysis of Tweets Using Naive Bayes*. Proceedings on 2016 2nd International Conference on Next Generation Computing Technologies, NGCT 2016, Dehradun, India, 14-16 Oktober 2016. 257–261.
- Han, J., Kamber, M., Pei, J., 2011. *Data Mining: Concepts and Techniques*. 3rd ed. Waltham: Morgan Kaufmann Publishers.
- Indriati, I., & Ridok, A., 2017. *Sentiment Analysis for Review Mobile Applications Using Neighbor Method Weighted K-Nearest Neighbor (Nwknn)*. Journal of Environmental Engineering and Sustainable Technology, 3(1), 23–32.
- Kikuchi, M., Yoshida, M., Okabe, M., & Umemura, K., 2015. *Confidence Interval of Probability Estimator of Laplace Smoothing*. ICAICTA 2015 - 2015 International Conference on Advanced Informatics: Concepts, Theory and Applications. Chonburi, Thailand, 19-22 Agustus 2015.
- Lestari, D, W, P., Perdana, R, S., Adikara, P, P., 2019. *Klasifikasi Video Clickbait pada Youtube Berdasarkan Analisis Sentimen Komentar Menggunakan Learning Vector Quantization (LVQ) dan Lexicon-Based Features*. Jurnal P Pengembangan Teknologi Informasi dan Ilmu Komputer, Vol.3, No.2, pp-1184-1189
- Liu, B., 2012. *Sentiment Analysis and Opinion Mining*. Synthesis Lectures on Human Language Technologies, 5(1), pp.1-167.
- Muljono, M., Artanti, D. P., Syukur, A., Prihandono, A., & Setiadi, D. R. I. M., 2018. *Analisa Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naïve Bayes*. Konferensi Nasional Sistem Informasi 2018. Pangkalpinang, Indonesia, 8-9 Maret 2018. pp165–170.
- Ratnawati, F., 2018. *Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter*. INOVTEK Polbeng - Seri Informatika, 3(1),
- Rofiqoh, U., Perdana, R. S., & Fauzi, M. A., 2017. *Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexicon Based Feature*. Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIHK) Universitas Brawijaya, 1(12), 1725–1732.
- Siddiqua, U. A., Ahsan, T., & Chy, A. N., 2017. *Combining a rule-based classifier with ensemble of feature sets and machine learning techniques for sentiment analysis on microblog*. 19th International Conference on Computer and Information Technology, ICCIT 2016. Dhaka, Bangladesh, 18-20 Desember 2016.
- Suhariyanto, Firmanto, A., & Sarno, R., 2018. *Prediction of Movie Sentiment based on Reviews and Score on Rotten Tomatoes using SentiWordnet*. Proceedings - 2018 International Seminar on Application for Technology of Information and Communication: Creative Technology for Human Life, ISemantic 2018. Semarang, Indonesia, 21-22 September 2018.
- Suryadi, Winarko, E., 2017. *Analisis Sentimen Review Hotel Menggunakan Algoritme Naïve Bayes Classifier dan Pendekatan Lexicon Based*. S2. Universitas Gadjah Mada. Tersedia melalui: Electronic Theses & Dissertations <http://etd.repository.ugm.ac.id/index.php?mod=penelitian_detail&sub=PenelitianDetail&act=view&typ=html&buku_id=111843&obyek_id=4>