Jurnal Informatika Universitas Pamulang **ISSN: 2541-1004**
Penerbit: Program Studi Teknik Informatika Universitas Pamulang **e-ISSN: 2622-4615**
Vol. 5, No. 3, September 2020 (295-301) **10.32493/informatika.v5i3.6110**

Klasifikasi *Rating* Otomatis pada Dokumen Teks Ulasan Produk Elektronik Menggunakan Metode *N-gram* dan *Naïve Bayes*

Rahmawan Bagus Trianto[1], Andri Triyono[2], Dhika Malita Puspita Arum[3]

[1,2,3]Ilmu Komputer, Fakultas Sains dan Kesehatan, Universitas An Nuur, Jl. Gajah Mada No.7, Majenang, Kuripan, Kec. Purwodadi, Kabupaten Grobogan, 58112
e-mail: [1]rahmawanbagust@gmail.com, [2]andritriyono1@gmail.com, [3]dhika.malita.11@gmail.com

Submitted Date: August 20th, 2020 Reviewed Date: September 22nd, 2020
Revised Date: September 28th, 2020 Accepted Date: September 30th, 2020

Abstract

Online product ratings usually provide descriptive reviews and also reviews in the form of ratings. Likewise, what was done at the Lazada online store. Descriptive review can provide a clear view compared to a rating review to other potential buyers. However, in reality there is a mismatch between the description review and the rating given. This creates a lack of information for sellers as well as potential buyers. Automatic classification of buyer descriptive reviews is proposed in this study so that there is a match between descriptive reviews and rating reviews. This automatic classification descriptive review uses the Naive Bayes algorithm with n-gram feature extraction and TF-IDF word weighting. The results of this study obtained the best accuracy of 94.06%, a recall of 91.73% and precision of 90.71% in Bigram feature extraction. With this accuracy value it can be used as a reference or model for classifying product description reviews, so that the feedback process between sellers and buyers can run well.

Keywords: Electronic product review; Classification; Naïve Bayes; n-gram; TF-IDF

Abstrak

Penilaian produk online biasanya dengan memberikan ulasan deskripsi dan juga ulasan berupa rating. Begitu juga yang dilakukan pada toko online Lazada. Ulasan deskripsi dapat memberikan pandangan yang jelas dibandingkan dengan ulasan berupa rating kepada calon pembeli lain. Namun pada kenyataannya sering dijumpai ketidaksesuaian antara ulasan deskripsi dengan rating yang diberikan. Hal ini membuat kurangnya informasi bagi penjual dan juga calon pembeli. Pengklasifikasian otomatis ulasan deskripsi pembeli diusulkan pada penelitian ini agar terjadi kesesuaian antara ulasan deskripsi dengan ulasan rating. Pengklasifikasian otomatis ulasan deskripsi ini menggunakan algoritma Naive Bayes dengan ekstraksi fitur n-gram dan pembobotan kata TF-IDF. Hasil dari penelitian ini didapatkan akurasi terbaik sebesar 94.06%, *recall* sebesar 91.73% dan presisi sebesar 90.71% pada ekstraksi fitur Bigram. Dengan nilai akurasi yang cukup tinggi tersebut dapat dijadikan salah satu acuan atau model untuk mengklasifikasikan ulasan deskripsi produk, sehingga proses umpan balik antara penjual dan pembeli dapat berjalan dengan baik.

Kata kunci: Ulasan produk elektronik; Klasifikasi; *Naïve Bayes*; *n-gram*; TF-IDF

1. Pendahuluan

Belanja online sudah dikenal banyak orang karena memiliki banyak keunggulan. Beberapa alasan orang berbelanja online adalah tidak perlu datang ke toko, harga relatif rendah, tidak perlu antri, tidak perlu bermacet-macetan di jalan (Harahap, 2018). Dan sekarang dengan adanya wabah covid-19 yang mengharuskan masyarakat harus menerapkan protokol kesehatan secara benar, maka belanja online menjadi salah satu pilihan terbaik (Hardilawati, 2020). Jual beli secara online sudah tentu memberikan pengalaman yang berbeda jika dibandingkan dengan cara konvensional. Perbedaan pengalaman berbelanja tersebut seperti tidak dapat melihat dan mencoba produk aslinya. Maka, untuk dapat menilai produk yang dijual secara online biasanya digunakan fitur ulasan. Dengan ulasan yang ditulis oleh pembeli maka akan sangat membantu bagi calon pembeli untuk memutuskan berbelanja. Selain itu, ulasan juga

sangat penting bagi penjual, karena dapat digunakan untuk meningkatkan kualitas penjualannya di masa yang akan datang (Farki, Baihaqi, & Wibawa, 2016).

Produk elektronik semakin diperlukan di era teknologi seperti sekarang, mulai dari komputer, *smartphone*, TV, peralatan internet, laptop, dan lain-lain. Banyak produk elektronik yang beredar luas di pasaran, terutama toko online seperti Lazada. Lazada merupakan salah satu *e-commerce* terbesar yang ada di Indonesia dengan berbagai produk yang dijualnya. Salah satu jenis produk yang dijual adalah barang elektronik. Banyak ulasan dari pembeli untuk menilai kualitas barang, kualitas pelayanan penjual sampai pada kualitas pengiriman barang. Namun, fakta di kenyataan menunjukkan bahwa tidak semua pembeli memberikan ulasannya, baik berupa *rating* maupun secara deskripsi. *Rating* merupakan bentuk dari tingkat kepuasan orang yang melakukan belanja berupa ulasan setelah membeli suatu produk. Hal ini mengakibatkan informasi terkait produk tidak dapat diterima secara utuh, sehingga dapat mempengaruhi calon pembeli untuk melakukan transaksi pembelian (Sapuhtra, Fauzi, & Rahayudi, 2019). Selain itu, pada kenyataan di lapangan tidak sedikit dijumpai adanya ulasan yang tidak sesuai antara deskripsi dan *rating* yang diberikan (Farki et al., 2016). Oleh sebab itu dibutuhkan sistem untuk dapat mengklasifikasikan ulasan pembeli ke dalam *rating* yang sesuai secara otomatis.

Pada penelitian sebelumnya, membahas mengenai klasifikasi untuk mendeteksi dua buah kelas ulasan, yaitu spam atau bukan spam, menggunakan *Naïve Bayes* dan ekstraksi fitur *n-gram*. Penggunaan ekstraksi fitur *n-gram* pada metode *Naïve Bayes* terbukti dapat meningkatkan akurasi hingga 80.44% (Setyaji, Zidny, Prabowo, & Hertantyo, 2018). Kemudian penelitian selanjutnya tentang klasifikasi teks pengaduan online menggunakan metode *n-gram* dan *Neighbor Weighted K-Nearest Neighbor* (WM K-NN), di mana penggunaan ekstraksi fitur *n-gram* yaitu unigram tidak memiliki pengaruh yang berarti. Dengan hasil rata-rata *f-measurement* sebesar 75.25%, sama dengan algortima WM K-NN saja tanpa menggunakan ekstraksi fitur *n-gram* (Prasanti, Fauzi, & Furqon, 2018). Penelitian lain tentang prediksi *rating* pada review produk kecantikan menggunakan metode *Semantic Orientation Calculator* dan Regresi Linier menggunakan ekstraksi fitur *n-gram*. Hasil penelitian ini menunjukkan ekstraksi firut *n-gram,* yaitu bigram dan trigram memiliki pengaruh dalam meningkatkan akurasi prediksi (Sapuhtra et al., 2019). Pada penelitian selanjutnya tentang prediksi *rating* otomatis menggunakan metode *Naïve Bayes* dan *n-gram* pada ulasan produk kecantikan. Pada penelitian ini menghasilkan nilai akurasi terbaik dengan menggunakan kombinasi *unigram* dan *bigram* pada metode *Naïve Bayes* dengan akurasi mencapai 97% (Pujadayanti, Fauzi, & Sari, 2018). Penelitian selanjutnya mengenai analisis pengaruh *stemmer* pada Bahasa Indonesia terhadap sentiment terjemahan ulasan film. Pada penelitian ini menyatakan bahwa *stemming* tidak memiliki pengaruh pada akurasi, bahkan cenderung mengurangi efisiensi dari analisis sentiment (Agastya, 2018).

Berdasarkan uraian di atas, maka peneliti mengusulkan sebuah penelitian dengan judul ***Klasifikasi Rating Otomatis pada Dokumen Teks Ulasan Produk Elektronik Menggunakan Metode N-gram dan Naïve Bayes***. Adapun ekstraksi fitur *n-gram* ini digunakan untuk meningkatkan akurasi klasifikasi pada metode *Naïve Bayes*, yaitu unigram, bigram, dan trigram. Penelitian ini juga tidak menggunakan proses *stemming* pada tahap *preprocessing* karena tidak memberikan pengingkatan yang signifikan terhadap performa. Tujuan dari penelitian ini adalah membuat model klasifikasi *rating* otomatis untuk memudahkan penjual dalam meningkatkan kualitas penjualan dan memberikan pertimbangan bagi calon pembeli di *e-commerce* untuk memutuskan dalam bertransaksi.

2. Metode Penelitian

2.1. Data

Penelitian ini menggunakan dataset terbuka yang dapat diakses oleh siapa saja. Dataset ini diambil dari alamat https://www.kaggle.com/grikomsn/lazada-indonesian-reviews. Dataset ini berkaitan dengan ulasan dari pembeli di *e-commerce* Lazada untuk jenis barang elektronik. Data yang diambil untuk penelitian ini sebanyak 10.000 baris dengan kolom *reviewContent*, yaitu berupa data teks komentar dan *rating*. *Rating* terdiri dari angka yang dimulai dari 1 sampai dengan 5. Kolom *rating* merupakan jenis data label dalam pembelajaran terarah atau biasa disebut dengan *supervised learning*. Data diambil secara acak untuk menghindari kemungkinan subyektifitas yang terjadi.

2.2. Metode yang Diusulkan

Gambar 1. *Flowchart* metode yang diusulkan

Tahap pertama pada proses klasifikasi *rating* ulasan dilakukan *pre-processing*. Pada tahap *pre-processing* ini terdiri dari tiga proses, yaitu *tokenize, transform case* dan *filter stopwords*. Setelah melalui tahap *pre-processing* dilanjutkan dengan ekstraksi fitur. Ekstraksi fitur pada penelitian ini menggunakan *n-gram*, yaitu *unigram, bigram* dan *trigram*. Selanjutnya masuk ke tahap pembobotan *term* atau pembobotan kata menggunakan TF-IDF. Setelah didapatkan bobot masing-masing kata, dilanjutkan pada proses klasifikasi menggunakan metode *Naïve Bayes* dan dilakukan validasi menggunakan *k-fold validation*. Validasi menggunakan *k-fold validation* dengan nilai *k*=10. Selanjutnya dilakukan evaluasi dengan pengukuran *F-Measure* untuk mengetahui akurasi, presisi dan juga *recall*.

2.3. Tahap *Pre-Processing*

Langkah awal yang dilakukan dalam memproses dokumen teks adalah *pre-processing*. Proses yang dilakukan seperti *tokenize*, yaitu memecah kalimat menjadi bentuk yang lebih sederhana, yaitu kata atau dikenal dengan nama *term* (García Adeva, Pikatza Atxa, Ubeda Carrillo, & Ansuategi Zengotitabengoa, 2014). *Transform case* atau mengubah huruf ke dalam bentuk huruf kecil juga dilakukan pada tahap *pre-processing* ini (García Adeva et al., 2014). Tahap yang tidak kalah penting adalah *filter stopwords,* yaitu membuang informasi yang tidak relevan, tidak penting dan tidak dibutuhkan dalam suatu dokumen teks (Sheela, 2018). Pada penelitian ini tidak menggunakan tahap *stemming* karena pada penelitian (Agastya, 2018) tidak menunjukkan adanya perbaikan hasil akhir, bahkan menambah beban pada komputer karena membutuhkan waktu dan sumber daya yang lebih banyak.

2.4. N-Gram

Dalam suatu dokumen teks, bahasa atau ungkapan tidak hanya tersusun dari kata-kata parsial. Namun untuk membentuk makna yang sebenarnya bahasa atau ungkapan terdiri dari gabungan dua kata atau lebih. Gabungan dua kata atau lebih ini dalam *text mining* dikenal dengan nama *n-gram* (Pujadayanti et al., 2018) (Lidya, Sitompul, & Efendi, 2015). Sebagai contoh penggunaan *n-gram* dengan *n*=2 pada kalimat *"Laptop sampai dengan selamat, pelayanan penjual memuaskan, pengiriman sangat cepat"* adalah *"laptop sampai", "sampai dengan", "dengan selamat", "pelayanan penjual", "penjual memuaskan", "pengiriman sangat", "sangat cepat"*. Jika menggunakan *n*=3 maka kalimat dipecah setiap tiga kata.

2.5. Pembobotan *Term* TF-IDF

Term Frequency – Inverse Document Frequency atau sering dikenal dengan TF-IDF adalah teknik gabungan dari TF dan IDF untuk memberikan bobot dari *term* (Haq & Budi, 2019). TF-IDF sangat terkenal digunakan untuk pemrosesan bahasa alami (Trstenjak, Mikac, & Donko, 2014). Selain itu TF-IDF juga sangat baik dalam menentukan *term* penting dalam dokumen teks (AL-Smadi, Jaradat, AL-Ayyoub, & Jararweh, 2017). Untuk mendapatkan bobot *term* dapat dilihat pada persamaan 1.

$$TF * IDF(d,t) = TF(d,t) * log\frac{N}{df(t)} \quad (1)$$

Untuk:

- TF * IDF (d,t) = bobot *term t* pada dokumen *d*
- TF(d,t) = frekuensi kemunculan *term* t pada dokumen *d*
- N = total seluruh dokumen
- df(t) = jumlah dokumen yang terdapat *term t*

2.6. Klasifikasi dengan *Naïve Bayes*

Klasifikasi telah dipakai dalam berbagai bidang, salah satunya pada konteks dokumen. Klasifikasi dokumen teks sudah dipakai seperti pengelompokan jenis dokumen, klasifikasi spam atau bukan spam pada pesan elektronik, menentukan sentiment positif dan negatif, serta masih banyak lagi (Di Nunzio, 2014). Salah satu metode yang banyak digunakan untuk klasifikasi adalah *Naïve Bayes*. Metode *Naïve Bayes* merupakan metode klasifikasi berdasarkan teorema probabilitas dan statistik *Bayes* (Haq & Budi, 2019) (Santoso et al., 2020). *Naïve Bayes* dapat memperkirakan peluang yang mungkin terjadi di masa depan dengan melihat data-data sebelumnya dengan ciri khas berupa independensi dari masing-masing kondisi. Rumus *Naïve Bayes* dapat dilihat secara matematisnya pada persamaan 2.

$$P(C|X) = \frac{P(C|X)*P(C)}{P(x)} \quad (2)$$

Untuk:

- x = data yang belum diketahui kelasnya
- c = data yang telah diketahui kelasnya
- p(c|x) = peluang hipotesis c berdasarkan kondisi x (*posterior probability*)
- p(x|c) = peluang hipotesis x berdasarkan kondisi c (*likelihood*)
- p(c) = peluang hipotesis c (*class prior probability*)
- p(x) = peluang x (*predictor prior probability*)

Sedangkan untuk klasifikasi dengan data yang berlanjut digunakan persamaan 3.

$$P(A_i|C_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} exp^{-\frac{(A_i-\mu_{ij})^2}{2(\sigma_{ij})^2}} \quad (3)$$

Untuk:

- P = peluang
- A_i = atribut ke i
- C = kelas yang akan diklasifikasikan
- σ = varian seluruh atribut
- μ = rata-rata seluruh atribut

Adapun pengujian pada penelitian ini menggunakan *k-fold validation* dengan nilai *k*=10, yang berarti data akan dibagi menjadi 10 bagian dan setiap 9 bagian akan dipakai sebagai data *learning* atau pelatihan, dan 1 bagian dipakai sebagai data *testing* atau pengujian (Saputri, Mahendra, & Adriani, 2019).

2.7. Validasi dan Pengukuran Performa

Validasi yang digunakan pada penelitian ini adalah *k-fold validation* dengan *k* = 10. Ini berarti dataset dipecah menjadi 10 bagian, 9 bagian digunakan untuk proses pelatihan dan 1 bagian digunakan untuk proses pengujian. Proses ini dilakukan sebanyak 10 kali (Wahono, Herman, & Ahmad, 2014). Tahap ini dapat dilihat pada tabel 1.

Tabel 1. Pembagian dataset pada *10-fold validation*

Validasi ke-n	Pembagian dataset									
1	■									
2		■								
3			■							
4				■						
5					■					
6						■				
7							■			
8								■		
9									■	
10										■

Tanda warna hitam menandakan dataset yang dipakai untuk data pelatihan. Sedangkan pada bagian yang lain menandakan dataset yang dipakai untuk data pelatihan.

Untuk mengetahui performa klasifikasi, diperlukan sebuah pengukuran, yaitu akurasi, presisi dan *recall* (Deolika, Kusrini, & Luthfi, 2019). Pengukuran ini sering dikenal dengan sebutan *confusion matrix. Confusion matrix* biasa dipakai pada proses pembelajaran terarah. Pada *confusion matrix*, data pada kolom mewakili data yang diharapkan, sedangkan data pada baris mewakili data yang diprediksi (Dhande & Patnaik, 2014). Berbeda dengan klasifikasi dengan 2 buah

kelas, untuk menghitung akurasi, presisi dan *recall* pada klasifikasi dengan kelas lebih dari 2 menggunakan rata-rata (Sokolova & Lapalme, 2009). Untuk menghitung akurasi, presisi dan *recall* dapat dilihat pada persamaan 4, 5 dan 6.

$$Akurasi = \frac{\sum_{i=1}^{c} \frac{TPi + TNi}{TPi + TNi + FPi + FNi}}{c} * 100\% \quad (4)$$

$$Presisi = \frac{\sum_{i=1}^{c} TPi}{\sum_{i=1}^{c} (FPi + TPi)} * 100\% \quad (5)$$

$$Recall = \frac{\sum_{i=1}^{c} TPi}{\sum_{i=1}^{c} (TPi + FNi)} * 100\% \quad (6)$$

Di mana:

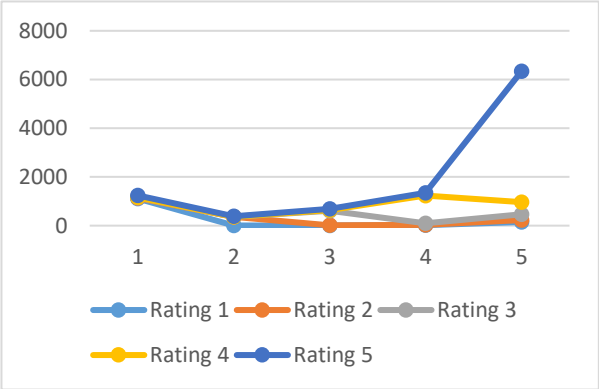
- c = jumlah kelas
- TPi = jumlah *True Positive* pada kelas ke i, jumlah data positif yang diklasifikasikan benar oleh sistem pada kelas ke i
- TNi = jumlah *True Negative* pada kelas ke i, jumlah data negatif yang diklasifikasikan benar oleh sistem pada kelas ke i
- FNi = jumlah *False Negative* pada kelas ke i, jumlah data negatif yang diklasifikasikan salah oleh sistem pada kelas ke i
- FPi = jumlah *False Positive* pada kelas ke i, jumlah data positif yang diklasifikasikan salah oleh sistem pada kelas ke i

3. Hasil dan Pembahasan

Tabel 2. Hasil pengujian klasifikasi

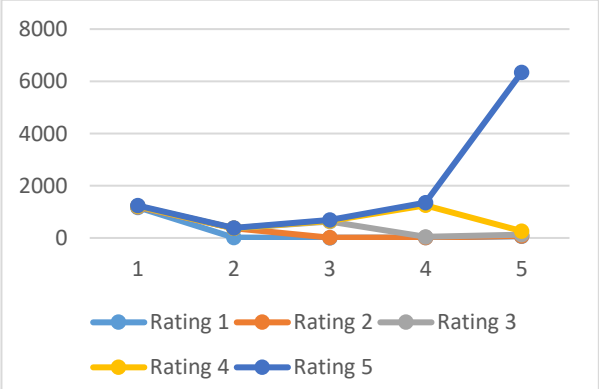
Model	Akurasi	Presisi	Recall
Naïve Bayes tanpa *n-gram*	85,66%	77.92%	87.28%
Naïve Bayes dengan *unigram*	85,66%	77.92%	87.28%
Naïve Bayes dengan *bigram*	**94,06%**	**90.71%**	**91.73%**
Naïve Bayes dengan *trigram*	89,91%	79.29%	91.48%

Tabel 2 menunjukkan hasil pengujian klasifikasi ulasan produk elektronik pada *e-commerce* Lazada. Hasil terbaik didapat pada pemakaian *bigram* pada metode *Naïve Bayes*, yaitu dengan nilai akurasi sebesar 94.06%, nilai presisi sebesar 90.71% dan nilai *recall* sebesar 91.73%. Dapat dilihat pula bahwa pemakaian *unigram* tidak memiliki pengaruh sama sekali pada hasil klasifikasi jika dibandingkan dengan metode *Naïve Bayes* saja, dibuktikan dengan nilai akurasi, presisi dan *recall* yang sama, yaitu masing-masing 85.66%, 77.92%, dan 87.28%. Hal ini karena penggunaan *unigram* pada dasarnya adalah sama saja dengan tahap *pre-processing*, yaitu pada tahap *tokenize* yang mana memecah kalimat menjadi kata-kata. Kemudian penggunaan *trigram* memberikan performa lebih baik jika dibandingkan dengan *Naïve Bayes* saja dan *Naïve Bayes* + *unigram*, namun lebih rendah jika dibandingkan dengan *Naïve Bayes* + *bigram*.



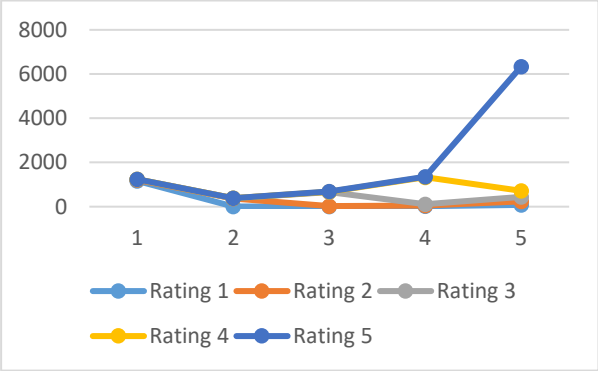
Gambar 2. Grafik sebaran klasifikasi *Naïve Bayes* dan *Naïve Bayes* + *Unigram*

Pada gambar 2 data sebaran *rating* pada hasil klasifikasi menggunakan *Naïve Bayes* dan *Naïve Bayes* + *Unigram* pada kelas *rating* 1 memiliki nilai yang benar sebanyak 1110 data, pada kelas *rating* 2 data yang benar sebanyak 347. Pada kelas *rating* 3 memiliki data yang benar sebanyak 603, pada kelas *rating* 4 memiliki data yang benar sebanyak 1139 buah, serta pada kelas *rating* 5 memiliki nilai yang benar sebanyak 5367 buah data.



Gambar 3. Grafik sebaran klasifikasi *Naïve Bayes* + *Bigram*

Pada gambar 3 dapat dilihat peningkatan hasil klasifikasi pada semua kelas *rating* jika dibandingkan dengan klasifikasi menggunakan *Naïve Bayes* saja maupun *Naïve Bayes + Unigram*, yaitu kelas *rating* 1 data yang benar sebanyak 1174, kelas *rating* 2 sebanyak 348, kelas *rating* 3 sebanyak 610, kelas *rating* 4 sebanyak 1203 dan kelas *rating* 5 sebanyak 6071.



Gambar 4. Grafik sebaran klasifikasi *Naïve Bayes + Trigram*

Gambar 4 menunjukkan hasil klasifikasi menggunakan *Naïve Bayes + Trigram*, di mana semua kelas *rating* menunjukkan hasil yang lebih baik jika dibandingkan dengan *Naïve Bayes* maupun *Naïve Bayes + Unigram*. Jika dibandingkan dengan *Naïve Bayes + Bigram,* hanya kelas *rating* 1 dan 5 yang memiliki data *true positive* lebih rendah, kelas *rating* 2, 3 dan 4 mempunyai data *true positive* lebih baik. Namun, pada kelas *rating* 2, 3 dan 4 tidak memiliki perbedaan yang besar. Pada kelas *rating* 5 memiliki perbedaan yang mencolok, dan ini salah satu sebab yang mengakibatkan akurasi, presisi dan *recall* menjadi lebih baik.

Pada pemakaian *n-gram*, terutama pada fitur *bigram*, dapat mengklasifikasikan lebih baik karena gabungan dua kata memberikan makna yang lebih baik, sehingga hasil klasifikasi juga menjadi lebih baik jika dibandingkan dengan gabungan dari tiga kata dan juga kata tunggal. Hal ini dapat dilihat seperti kalimat "*Barang Ok, kualitas bagus dengan harga yg cukup miring, kualitas gambar juga ok, pengiriman lumayan, ga nyesel pokonya beli TV ini*", setelah melalui tahap *pre-processing* dan ekstraksi fitur maka menjadi *"barang ok", "ok kualitas", "kualitas bagus", "bagus harga", "harga cukup", "cukup miring", "miring kualitas", "kualitas gambar", "gambar ok", "ok pengiriman", "pengiriman lumayan", "lumayan ga", "ga nyesel", "nyesel pokonya", "pokonya beli", "beli tv"*. Dari gambar 2, 3 dan 4 dapat dilihat pula bahwa sebaran dataset yang diambil memiliki sifat *imbalance dataset*, yaitu ketidakmerataan jumlah data antar kelas, terutama pada kelas *rating* 5.

4. Kesimpulan

Klasifikasi ulasan produk elektronik ke dalam bentuk *rating* dapat dilakukan menggunakan metode *Naïve Bayes* dan *n-gram*. Diawali dengan tahap *pre-processing* kemudian dilanjutkan pembobotan *term* menggunakan TF-IDF serta ekstraksi fitur menggunakan *n-gram* dan dilanjutkan klasifikasi menggunakan metode *Naïve Bayes*. Ekstraksi fitur menggunakan *n-gram* terbukti dapat meningkatkan hasil akhir klasifikasi. Hasil terbaik yang didapatkan adalah menggunakan ekstraksi fitur *bigram*, diikuti ekstraksi fitur *trigram* dan setelahnya ekstraksi fitur *unigram* (dan juga hanya *Naïve Bayes* tanpa *n-gram*). Adapun hasil akurasi, presisi dan *recall* pada klasifikasi dengan *Naïve Bayes* menggunakan ekstraksi fitur *bigram* adalah sebesar 94.06%, 90.71%, dan 91.73%. Dengan nilai akurasi yang cukup tinggi tersebut dapat dijadikan salah satu acuan atau model untuk mengklasifikasikan ulasan deskripsi produk, sehingga proses umpan balik antara penjual dan pembeli dapat berjalan dengan baik. Untuk hasil akurasi, presisi dan *recall* pada klasifikasi dengan *Naïve Bayes* menggunakan ekstraksi fitur *trigram* sebesar 89.91%, 79.29% dan 91.48%. Kemudian hasil akurasi, presisi dan *recall* pada klasifikasi dengan *Naïve Bayes* menggunakan ekstraksi fitur *unigram* dan juga tanpa *n-gram* memiliki nilai yang sama, yaitu 85.66%, 77.92% dan 87.28%.

Referensi

Agastya, I. M. A. (2018). Pengaruh Stemmer Bahasa Indonesia Terhadap Peforma Analisis Sentimen Terjemahan Ulasan Film. *Jurnal Tekno Kompak*, *12*(1), 18. https://doi.org/10.33365/jtk.v12i1.70

AL-Smadi, M., Jaradat, Z., AL-Ayyoub, M., & Jararweh, Y. (2017). Paraphrase identification and semantic text similarity analysis in Arabic news tweets using lexical, syntactic, and semantic features. *Information Processing & Management*, *53*(3), 640–652. https://doi.org/10.1016/j.ipm.2017.01.002

Deolika, A., Kusrini, K., & Luthfi, E. T. (2019). Analisis Pembobotan Kata Pada Klasifikasi Text Mining. *Jurnal Teknologi Informasi*, *3*(2), 179. https://doi.org/10.36294/jurti.v3i2.1077

Dhande, L. L., & Patnaik, P. G. K. (2014). Analyzing Sentiment of Movie Review Data using Naive Bayes Neural Classifier. *International Journal of*

Emerging Trends & Technology in Computer Science (IJETTCS), *3*(4), 313–320. Diambil dari www.ijettcs.org

Di Nunzio, G. M. (2014). A new decision to take for cost-sensitive Naïve Bayes classifiers. *Information Processing & Management*, *50*(5), 653–674. https://doi.org/10.1016/j.ipm.2014.04.008

Farki, A., Baihaqi, I., & Wibawa, M. (2016). Pengaruh online customer review rating terhadap kepercayaan place di indonesia. *Jurnal Teknik ITS*, *5*(2), A614–A619.

García Adeva, J. J., Pikatza Atxa, J. M., Ubeda Carrillo, M., & Ansuategi Zengotitabengoa, E. (2014). Automatic text classification to support systematic reviews in medicine. *Expert Systems with Applications*, *41*(4), 1498–1508. https://doi.org/10.1016/j.eswa.2013.08.047

Haq, F. I. N., & Budi, E. (2019). Implementasi Naive Bayes Classifier untuk Prediksi Kepribadian Big Five pada Twitter Menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) dan Term Frequency-Relevance Frequency (TF-RF) Program Studi Sarjana Ilmu Komputasi Fakultas Informatik. *e-Proceeding of Engineering*, *6*(2), 9785–9795.

Harahap, D. A. (2018). Perilaku Belanja Online Di Indonesia: Studi Kasus. *JRMSI - Jurnal Riset Manajemen Sains Indonesia*, *9*(2), 193–213. https://doi.org/10.21009/jrmsi.009.2.02

Hardilawati, W. L. (2020). Jurnal Akuntansi & Ekonomika. *Jurnal Akuntansi & Ekonomika*, *10*(1), 89–98. Diambil dari http://ejurnal.umri.ac.id/index.php/jae

Lidya, S. K., Sitompul, O. S., & Efendi, S. (2015). Sentiment Analysis Pada Teks Bahasa Indonesia Menggunakan Support Vector Machine (Svm). *Seminar Nasional Teknologi dan Komunikasi 2015*, *2015*(Sentika), 1–8.

Prasanti, A. A., Fauzi, M. A., & Furqon, M. T. (2018). Klasifikasi Teks Pengaduan Pada Sambat Online Menggunakan Metode N- Gram dan Neighbor Weighted K-Nearest Neighbor (NW-KNN). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, *2*(2), 594–601.

Pujadayanti, I., Fauzi, M. A., & Sari, Y. A. (2018). Prediksi Rating Otomatis pada Ulasan Produk Kecantikan dengan Metode Naïve Bayes dan N-gram. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK)*, *2*(11), 4421–4427.

Santoso, H. A., Rachmawanto, E. H., Nugraha, A., Nugroho, A. A., Setiadi, D. R. I. M., & Basuki, R. S. (2020). Hoax classification and sentiment analysis of Indonesian news using Naive Bayes optimization. *Telkomnika (Telecommunication Computing Electronics and Control)*, *18*(2), 799–806. https://doi.org/10.12928/TELKOMNIKA.V18I2.14744

Sapuhtra, B. D., Fauzi, M. A., & Rahayudi, B. (2019). Prediksi Rating Pada Review Produk Kecantikan Menggunakan Metode Semantic Orientation Calculator dan Regresi Linier. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, *3*(5), 4477–4483.

Saputri, M. S., Mahendra, R., & Adriani, M. (2019). Emotion Classification on Indonesian Twitter Dataset. *Proceedings of the 2018 International Conference on Asian Language Processing, IALP 2018*, 90–95. IEEE. https://doi.org/10.1109/IALP.2018.8629262

Setyaji, M., Zidny, M., Prabowo, W. A., & Hertantyo, G. B. (2018). Naive Bayes dengan Ekstraksi Fitur N-gram dalam Mendeteksi Spam Ulasan Bahasa Indonesia. *Proceedings on Conference on Electrical Engineering, Telematics, Industrial Technology, and Creative Media*, 56–60.

Sheela, S. P. (2018). Sentiment Analysis and Prediction of Online Reviews with Empty Ratings. *International Journal of Applied Engineering Research*, *13*(14), 11532–11539.

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, *45*(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Trstenjak, B., Mikac, S., & Donko, D. (2014). KNN with TF-IDF based framework for text categorization. *Procedia Engineering*, *69*, 1356–1364. Elsevier B.V. https://doi.org/10.1016/j.proeng.2014.03.129

Wahono, R. S., Herman, N. S., & Ahmad, S. (2014). A comparison framework of classification models for software defect prediction. *Advanced Science Letters*, *20*(10–12), 1945–1950. https://doi.org/10.1166/asl.2014.5640