

QUERY EXPANSION RANKING PADA ANALISIS SENTIMEN MENGGUNAKAN KLASIFIKASI *MULTINOMIAL NAÏVE BAYES*

(Studi Kasus : Ulasan Aplikasi Shopee pada Hari Belanja *Online* Nasional 2020)

Lutfiah Maharani Siniwi¹, Alan Prahutama², Arief Rachman Hakim³

^{1,2,3} Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro
lutfiahmaharani27@gmail.com

ABSTRACT

Shopee is one of the e-commerce sites that has many users in Indonesia. Shopee provides various attractive promos on special days such as National Online Shopping Day on December 12. Shopee site was a complete error on December 12, 2020. Complaints and opinions of Shopee users were also shared through various media, one of them was Google Play Store. Sentiment analysis was used to see the user's response to the Shopee's incident. Sentiment analysis results can be extracted to obtain information regarding positive or negative reviews from Shopee users. Sentiment analysis was performed using the Multinomial Naïve Bayes classification, the simplest method of probability classification, but it is sensitive to feature selection so that the amount of data is determined by the results of feature selection Query Expansion Ranking. The algorithm that has the highest accuracy and kappa statistic is the best algorithm in classifying Shopee's users sentiment. The results showed that the classification performance using Multinomial Naïve Bayes with 80% of the features (terms) which have the highest Query Expansion Ranking value was obtained at the accuracy and kappa statistics values are 89% and 77.62%. This means that Multinomial Nave Bayes has a good performance in classifying reviews and the number of features used affects the performance results obtained.

Keywords: Shopee, National Online Shopping Day, Multinomial Naïve Bayes, Query Expansion Ranking

1. PENDAHULUAN

Gaya hidup dalam memenuhi kebutuhan dengan membeli secara langsung mulai ditinggalkan sebagai pencegahan untuk mengurangi persebaran Covid-19 dengan memanfaatkan jasa jual beli *online* melalui *e-commerce*. Platform *e-commerce* yang paling banyak digunakan oleh masyarakat Indonesia adalah *marketplace*. Shopee merupakan aplikasi *mobile marketplace* yang memiliki jumlah pengunjung tertinggi sebanyak 129 juta per bulan (iPrice, 2020). Shopee telah melakukan perubahan untuk menarik minat pelanggan agar lebih banyak bertransaksi melalui situs tersebut, seperti pemberian gratis ongkos kirim, *cashback* bahkan diskon besar-besaran. Pemberian berbagai promo tersebut akan lebih menarik lagi pada hari-hari spesial seperti Hari Belanja *Online* Nasional (Harbolnas) pada tanggal 12 Desember. Situs Shopee mengalami *error* secara menyeluruh pada tanggal 12 Desember 2020. Keluhan dan opini pengguna Shopee ini pun disalurkan melalui berbagai sarana salah satunya layanan Google Play Store. Analisis sentimen dilakukan untuk melihat respon pengguna mengenai aplikasi Shopee pada kejadian tersebut

Pada penelitian ini dilakukan klasifikasi ulasan pengguna Shopee ke dalam dua sentimen, yaitu sentimen positif dan sentimen negatif. Pengklasifikasian dilakukan dengan metode *Multinomial Naïve Bayes*. Klasifikasi *Multinomial Naïve Bayes* merupakan pengembangan model dari algoritma bayes yang biasa digunakan dalam klasifikasi teks. Model multinomial memperhitungkan frekuensi dari setiap kata yang muncul pada dokumen di mana perhitungan hasil frekuensi ini dapat membantu klasifikasi (McCallum & Nigam, 1998). Klasifikasi *Naïve Bayes* masih memiliki beberapa kekurangan yakni sangat sensitif dalam pemilihan fitur (Chen, Huang, Tian, & Qu, 2009). Pemilihan fitur dapat membuat pengklasifikasi baik lebih efisien atau efektif dengan mengurangi jumlah data yang dianalisis, salah satunya dengan *Query Expansion Ranking*. *Query Expansion Ranking* adalah sebuah metode seleksi fitur yang berguna untuk mengurangi kompleksitas komputasi tanpa mengurangi kualitas dari analisis sentimen (Fauzi, Arifin, & Gosaria, 2017). Hasil pengklasifikasian sentimen yang diperoleh dapat menjadi acuan dalam pengembangan *e-commerce* kedepannya. Berdasarkan penjelasan tersebut, tujuan yang hendak dicapai adalah mendapatkan nilai akurasi dan *kappa statistic* hasil kinerja klasifikasi sentimen terkait ulasan aplikasi Shopee menggunakan klasifikasi *Multinomial Naïve Bayes* dengan seleksi fitur *Query Expansion Ranking*. Melalui ulasan yang ditulis pengguna aplikasi Shopee diharapkan menjadi masukan guna perbaikan untuk waktu yang akan datang.

2. TINJAUAN PUSTAKA

2.1. E-commerce Shopee

Shopee Indonesia merupakan salah satu *marketplace* yang dimiliki oleh *SEA Group* yang terdapat di Singapura. Bisnis *C2C (Customer to Customer) mobile marketplace* yang dikelola Shopee memungkinkan untuk mampu diterima dengan mudah oleh berbagai lapisan masyarakat, termasuk di Indonesia. Sejak peluncurannya, Shopee Indonesia mengalami perkembangan yang sangat pesat, bahkan hingga tahun 2020 aplikasinya sudah di-*download* oleh lebih dari 129 juta pengguna (iPrice, 2020). Sasaran pengguna Shopee adalah kalangan muda yang saat ini terbiasa melakukan kegiatan dengan bantuan *smartphone* termasuk kegiatan berbelanja aplikasi mobile guna untuk menunjang kegiatan berbelanja yang mudah dan cepat.

2.2. Hari Belanja Online Nasional

Hari Belanja *Online Nasional* (Harbolnas) merupakan kegiatan tahunan yang diprakarsai bersama oleh enam *e-commerce* besar di Indonesia sejak 12 Desember 2012. Enam *e-commerce* tersebut yaitu Lazada Indonesia, Zalora, Blanja, PinkEmma, Berrybenka, dan Bukalapak. Harbolnas 2020 yang diawasi oleh Asosiasi *E-Commerce* Indonesia (idEA) berkembang signifikan terbukti dengan jumlah pesertanya yang mencapai lebih dari 250 *platform* belanja *online* (idEA, 2020).

2.3. Text Mining

Menurut Feldman dan Sanger (2007), text mining adalah proses mengekstrak informasi yang berguna dari sekumpulan dokumen dari waktu ke waktu melalui identifikasi pola yang menarik. Text mining adalah proses penemuan informasi baru dengan mengekstrak pola secara otomatis dari berbagai sumber teks (Kini, et al., 2015).

2.4. Analisis Sentimen

Analisis sentimen adalah bidang studi yang menganalisis opini, sentimen, evaluasi, penilaian, sikap, dan perasaan terhadap entitas seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik, dan atributnya (Liu, 2012). Ling, et al. (2014) menyatakan bahwa tugas dasar dalam analisis sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, kalimat, atau pendapat. Polaritas mempunyai arti bahwa teks yang ada dalam dokumen, kalimat, atau pendapat memiliki aspek positif atau negatif.

2.5. Pre-Processing Data

Pre-Processing data sangat penting dalam melakukan analisis sentimen, terutama untuk media sosial yang sebagian besar berisi kata-kata atau kalimat yang tidak formal dan tidak terstruktur serta memiliki *noise* yang besar (Mujilahwati, 2016). Tahapan *pre-processing* yang akan dilakukan adalah sebagai berikut: *case folding*, *cleaning*, dan normalisasi kata.

2.6. Sentiment Scoring

Sentiment scoring dilakukan dengan menggunakan kamus sentimen, *boosterwords*, dan negasi. Kamus sentimen berisi kumpulan kata yang telah diberi kekuatan sentimen, yaitu kekuatan negatif antara -1 hingga -5 dan kekuatan positif antara +1 hingga +5. *Boosterwords* adalah kata yang dapat meningkatkan atau mengurangi intensitas sentimen kata di sebelahnya. Kata negasi merupakan kata yang terdapat dalam suatu kalimat yang dapat mengubah orientasi dari suatu opini (Wahid & Azhari, 2016).

2.7. Seleksi Fitur

Seleksi fitur merupakan tahapan mengurangi jumlah fitur yang menentukan suatu nilai kelas target, fitur irelevan, berlebihan dan data yang menyebabkan salah memberi arti pada kelas target, dan berdampak langsung pada aplikasi (Maulida, Suyatno, & Hatta, 2016). Tahapan seleksi fitur adalah *stopwords removal*, *stemming*, *tokenizing*, dan *Query Expansion Ranking*. *Stopwords removal* dilakukan untuk menghilangkan kata yang sering muncul tetapi tidak relevan. Tahap *stemming* merupakan perubahan kata imbuhan menjadi kata dasar. Tahap *tokenizing* merupakan proses pembagian kalimat menjadi potongan kata-kata.

Query Expansion Ranking (QER) adalah sebuah metode seleksi fitur yang terinspirasi dari metode *Query Expansion* yang berguna untuk meningkatkan kualitas *query* yang dimasukkan oleh pengguna kemudian digabung dengan cara *probabilistic weighting model* untuk memberi skor pada setiap fitur (Parlar & Özel, 2016). Perhitungan dari nilai QER diperoleh persamaan:

$$score_f = \frac{|p_f + q_f|}{|p_f - q_f|} \quad (1)$$

p_f adalah peluang *term f* dalam dokumen kelas positif dan q_f dan peluang *term f* dalam dokumen kelas negatif. Nilai-nilai peluang tersebut dihitung berdasarkan perhitungan:

$$p_f = \frac{df_+^f + 0,5}{n^+ + 1,0} \quad (2)$$

$$q_f = \frac{df_-^f + 0,5}{n^- + 0,5} \quad (3)$$

df_+^f dan df_-^f adalah jumlah dokumen positif dan negatif yang mengandung *term* f , sedangkan n^+ dan n^- adalah jumlah dokumen positif dan negatif.

2.8. Pembobotan Kata

Pembobotan kata yang digunakan ialah metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Manning, et al. (2009) menyatakan bahwa *Term Frequency* merupakan banyaknya jumlah kata atau *term* tertentu yang ada dalam suatu dokumen. Sementara *Inverse Document Frequency* adalah frekuensi kemunculan kata atau *term* pada seluruh dokumen. Persamaan untuk *term frequency-inverse document frequency* (TF-IDF) dapat dihitung menggunakan rumus sebagai berikut:

$$W_{j,i} = \frac{n_{j,i}}{\sum_{c_k} n_{c_k,i}} \cdot \log_2 \frac{D}{d_j} \quad (4)$$

$W_{j,i}$ adalah pembobotan *TF-IDF* untuk *term* ke- j pada dokumen ke- i , $n_{j,i}$ merupakan jumlah kemunculan *term* ke- j pada dokumen ke- i , $\sum_{c_k} n_{c_k,i}$ jumlah kemunculan seluruh *term* pada dokumen ke- i pada kelas k , D adalah banyaknya dokumen yang dibangkitkan, dan d_j merupakan banyaknya dokumen yang mengandung *term* ke- j .

2.9. Klasifikasi *Multinomial Naïve Bayes*

Klasifikasi *Multinomial Naïve Bayes* merupakan pengembangan model dari algoritma bayes yang biasa digunakan dalam klasifikasi teks. Dokumen pada *Multinomial Naïve Bayes* dianggap sebagai “*bag of words*” di mana urutan kejadian munculnya kata dalam dokumen diabaikan, sehingga setiap kata diproses menggunakan distribusi multinomial. Model multinomial memperhitungkan frekuensi dari setiap kata yang muncul pada dokumen di mana perhitungan hasil frekuensi ini dapat membantu klasifikasi (McCallum & Nigam, 1998). Klasifikasi dokumen bertujuan untuk menentukan kelas terbaik untuk suatu dokumen. Kelas terbaik dalam klasifikasi *Naïve Bayes* ditentukan dengan mencari *maximum a posteriori* (MAP) kelas C_{MAP} melalui persamaan:

$$C_{MAP} = \arg \max_{c \in C} \hat{P}(c) \prod_{j=1}^n \hat{P}(f_j | c_k) \quad (5)$$

$\hat{P}(c)$ dan $\hat{P}(f_j | c_k)$ dihitung pada saat *training*. Persamaannya dapat dituliskan sebagai berikut:

$$\hat{P}(c_k) = \frac{N^{c_k}}{N} \quad (6)$$

$$\hat{P}(f_j | c_k) = \frac{w_{c_k f_j} + 1}{(\sum_{w' \in V} w_{c_k f_j'}) + B'} \quad (7)$$

N^{c_k} adalah jumlah dokumen *training* dalam kelas k , N adalah jumlah keseluruhan dokumen *training* dari seluruh kelas, B' merupakan jumlah IDF seluruh *term* pada *vocabulary*, $w_{c_k f_j}$ adalah bobot TF-IDF *term* f_j pada dokumen dengan kelas k , dan $\sum_{w' \in V} w_{c_k f_j'}$ adalah jumlah bobot TF-IDF seluruh *term* pada kelas k .

2.10. Evaluasi Performansi

Proses evaluasi akurasi klasifikasi dilakukan dengan memperhatikan *confusion matrix* (Said & Pranoto, 2015). *Confusion matrix* adalah suatu metode yang biasanya digunakan untuk melakukan perhitungan akurasi pada konsep *data mining*. Berikut merupakan tabel *confusion matrix*.

Tabel 1. Confusion Matrix

<i>Corect Classification</i>	<i>Classified as</i>	
	<i>Predicted “+”</i>	<i>Predicted “-“</i>
<i>Actual “+”</i>	<i>True Positives</i>	<i>False Negatives</i>
<i>Actual “-“</i>	<i>False Positives</i>	<i>True Negatives</i>

Ukuran yang dapat digunakan untuk menilai atau mengevaluasi model klasifikasi yaitu akurasi dan *kappa statistics*. Akurasi merupakan nilai proporsi dari jumlah tiap kelas yang diklasifikasikan dengan benar terhadap total semua data (Han, Kamber, & Pei, 2012). Sedangkan, *Kappa Statistic* adalah analisis statistik berdasarkan kesepakatan penafsir untuk data kualitatif (Vierra & Garrett, 2015). Keduanya dapat dihitung dengan persamaan :

$$Accuracy = P_0 = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$P_c = \left[\left(\frac{TP + FP}{Total} \right) \left(\frac{TP + FN}{Total} \right) \right] + \left[\left(\frac{FN + TN}{Total} \right) \left(\frac{FP + TN}{Total} \right) \right] \quad (9)$$

$$K = \frac{P_0 - P_c}{(1 - P_c)} \quad (10)$$

2.11. Text Visualisation

Text visualisation merupakan teknik yang digunakan untuk menampilkan visualisasi data teks. Salah satu bentuk dari visualisasi teks adalah *word cloud*. Word cloud memberikan informasi yang relevan dengan menunjukkan istilah atau *term* yang penting dalam teks (Tessem, Bjørnstad, Chen, & Nyre, 2015).

3. METODOLOGI PENELITIAN

3.1. Jenis dan Sumber Data

Data yang digunakan pada penelitian ini adalah data kualitatif berupa ulasan pada aplikasi Shopee di Google Play Store. Data yang akan diambil melalui *web scraping* pada penelitian ini berupa tanggal, nama, ulasan pengguna, dan *rating* dari aplikasi Shopee di Google Play Store pada laman <https://play.google.com/store/apps/details?id=com.shopee.id>. Data ulasan bahasa Indonesia yang diambil adalah data ulasan aplikasi Shopee pada tanggal 9 – 15 Desember 2020 dan diperoleh sebanyak 2439 data ulasan, dari 2439 ulasan tersebut dipilih sebanyak 1000 ulasan dalam kategori bahasa Indonesia dengan perbandingan data training dan data testing sebesar 80%:20%.

3.2. Langkah-Langkah Analisis Data

Penelitian ini akan dilakukan dengan bantuan software Data Miner untuk web scraping, RStudio, dan Microsoft Excel untuk analisis data. Adapun tahapan-tahapan analisis yang digunakan dalam penelitian ini diuraikan sebagai berikut:

1. *Scraping data*

Tahap pertama adalah mengambil seluruh data sebesar 2439 ulasan, kemudian dipilih sebanyak 1000 ulasan dalam kategori bahasa Indonesia.

2. *Pre-Processing data*

Sebelum digunakan perlu melakukan kan *pre-processing* pada data. *Pre-processing* yang dilakukan antara lain adalah *case folding*, *cleaning*, dan normalisasi kata.

3. Pelabelan data dengan *sentiment scoring* dan manual
Melakukan pelabelan data menjadi label positif atau negatif dengan menggunakan *sentiment scoring* lalu dicek secara manual.
4. *Stopwords removal*
Data yang telah melalui tahap pelabelan selanjutnya dihilangkan kata-kata yang tidak memiliki makna sesuai dengan *stoplist* atau daftar kata yang dapat dihilangkan.
5. *Stemming* dan *tokenizing*
Tahap selanjutnya adalah *stemming* dan *tokenizing*. Kata-kata yang menyusun dokumen tersebut dipisahkan sehingga dapat dilakukan *stemming*, yakni proses pemotongan imbuhan atau pengembalian kata berimbuhan menjadi kata dasar.
6. Pembobotan kata
Metode yang digunakan pada pembobotan kata adalah pembobotan dengan TF-IDF. Hasil pembobotan akan digunakan untuk proses klasifikasi.
7. Seleksi fitur *Query Expansion Ranking*
Proses menghilangkan fitur yang kurang relevan untuk mengurangi kompleksitas komputasi dengan *Query Expansion Ranking* dengan menggunakan presentase fitur sebesar 100%, 80%, 60%, 40%, dan 20%.
8. Membangun model klasifikasi dan pengujian dengan *Multinomial Naïve Bayes*
Setelah melakukan seleksi fitur, akan dilakukan klasifikasi pada dokumen yang diperlukan dengan metode *Multinomial Naïve Bayes*.
9. Evaluasi performansi
Evaluasi performansi menggunakan *confusion matrix*.
10. *Text visualisation*
Teknik yang digunakan dalam tahap ini adalah *word cloud* dengan memanfaatkan frekuensi kata pada seluruh dokumen.

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Pengumpulan data pada aplikasi Shopee di Google Play Store menggunakan metode *scraping data* dengan menggunakan aplikasi Data Miner dengan total ulasan yang berhasil dikumpulkan adalah sebanyak 2439. Data yang diambil berupa tanggal, nama, ulasan pengguna, dan *rating* dari aplikasi Shopee. Ulasan yang akan diolah adalah 1000 ulasan pada Hari Belanja Online Nasional (9 – 15 Desember 2020) dengan kategori bahasa Indonesia.

4.2. *Pre-Processing Data*

Proses ini dilakukan untuk mengatasi kecenderungan data yang tidak terstruktur. Berikut merupakan tahapan *pre-processing* ulasan:

1. *Case folding*: Merubah keseluruhan teks menjadi huruf kecil agar seragam.
2. *Cleaning*: Menghapus karakter-karakter yang tidak dibutuhkan seperti huruf atau karakter di luar dari alfabet a-z, angka, tanda baca, dan *emoticon*.
3. Normalisasi kata: Mentransformasi kata-kata yang tidak sesuai Kamus Besar Bahasa Indonesia (KBBI) dan memperbaiki penulisan suatu kata.

4.3. Pelabelan Data

Berdasarkan pelabelan yang telah dilakukan berdasarkan skor sentimen menggunakan kamus sentimen, *boosterwords*, dan negasi diperoleh 558 ulasan berlabel positif dan 442 ulasan berlabel negatif. Namun, terdapat 44 ulasan yang salah dalam pelabelan dari 1000 data ulasan. Hal tersebut mengartikan bahwa sebesar 4,4% ulasan salah dalam pelabelan data menggunakan ketiga kamus *sentiment scoring*. Setelah dilakukan perbaikan pada ulasan yang salah dalam pelabelan, jumlah ulasan positif menjadi sebanyak 514 dan ulasan negatif sebanyak 486. Hal ini disebabkan oleh hasil skor akhir yang dihitung oleh program berbeda dengan perhitungan manual.

4.4. Stopwords Removal

Tahap *stopwords removal* menggunakan kamus *stoplist* yang memuat 771 kata. Setelah melalui *stopwords removal*, jumlah kata pada seluruh dokumen adalah sebanyak 1455 fitur (*term*) yang sebelumnya berjumlah sebanyak 1759.

4.5. Stemming dan Tokenizing

Tahap *stemming* menggunakan *package* *katadasaR* pada Rstudio untuk mengubah kata imbuhan menjadi kata dasar. Tahap *tokenizing* memisahkan dokumen menjadi potongan kata dengan acuan pemisah berupa spasi.

4.6. Pembobotan Kata

Pembobotan kata menggunakan perhitungan *Term Frequency-Inverse Document Frequency* (TF-IDF). Berikut hasil nilai pembobotan kata menggunakan TF-IDF yang dapat dilihat pada Tabel 2.

Tabel 2. Pembobotan Kata dengan TF-IDF

Ulasan ke-	adakan	andal	gratis	kirim	ongkos	...	untung
286	0	0	1,09521	0,81288	1,74145	...	0
629	0	0,73808	0,43808	0,32515	0	...	0
771	0,32975	0	0	0,17113	0,73324	...	0
869	0	0	0,43808	0,32515	0	...	0,67959
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1000	0	0	0	0	0	...	0

4.7. Klasifikasi

4.7.1. Data Training dan Data Testing

Perbandingan data *training* dan data *testing* yang digunakan dalam penelitian ini adalah 80% : 20%. Tabel 3 merupakan perbandingan proporsi data *training* dan data *testing*.

Tabel 3. Proporsi Data Training dan Data Testing

Klasifikasi	Data Training	Data Testing	Total
Positif	411	103	514
Negatif	389	97	486
Total	800	200	1000

4.7.2. Query Expansion Ranking

Proses seleksi fitur *query expansion ranking* (QER) ini adalah proses untuk mendapatkan nilai fitur pada data *training*. Hasil persamaan tersebut kemudian diurutkan dari nilai terkecil hingga terbesar. Berikut hasil perhitungan QER yang dapat dilihat pada Tabel 4.

Tabel 4. Hasil Perhitungan Nilai QER

Ranking	Fitur	df_+^f	p_f	df_-^f	q_f	$score_f$
1	out	1	0,003641	91	0,234917	1,031484
2	error	1	0,003641	87	0,224647	1,032947
3	check	1	0,003641	81	0,209243	1,035416
4	ganggu	0	0,001214	25	0,065469	1,037774
5	masuk	0	0,001214	21	0,055199	1,04496
6	robot	0	0,001214	17	0,044929	1,055522
7	hilang	0	0,001214	15	0,039795	1,062911
7	log	0	0,001214	15	0,039795	1,062911
⋮	⋮	⋮	⋮	⋮	⋮	⋮
140	toko	9	0,023058	8	0,021823	36,32913

4.7.3. Multinomial Naïve Bayes

Confusion matrix adalah evaluasi performansi yang digunakan untuk menunjukkan parameter yang tepat untuk digunakan dalam mengukur performa. Berikut merupakan hasil dari klasifikasi *Multinomial Naïve Bayes* dengan berbagai jumlah fitur (term) yang digunakan.

1. Penggunaan 100% Fitur

Seluruh fitur (*term*) yang ada pada data *test* digunakan pada proses klasifikasi pertama. Jumlah fitur yang digunakan sebanyak 1110 fitur.

Tabel 5. *Confusion Matrix* dengan 100% Fitur

Kelas Prediksi	Kelas Aktual	
	Negatif	Positif
Negatif	72	12
Positif	15	101

2. Penggunaan 80% Fitur

Proses klasifikasi kedua dilakukan dengan 80% fitur (*term*) yang memiliki peringkat tertinggi pada hasil pengolahan nilai QER data *training*. Jumlah fitur yang digunakan sebanyak 946 fitur.

Tabel 6. *Confusion Matrix* dengan 80% Fitur

Kelas Prediksi	Kelas Aktual	
	Negatif	Positif
Negatif	76	11
Positif	11	102

3. Penggunaan 60% Fitur

Pada proses klasifikasi berikutnya, terdapat 60% fitur (*term*) peringkat QER tertinggi pada data *training* digunakan sebanyak 845 fitur.

Tabel 7. *Confusion Matrix* dengan 60% Fitur

Kelas Prediksi	Kelas Aktual	
	Negatif	Positif
Negatif	69	11
Positif	18	101

4. Penggunaan 40% Fitur

Terdapat 40% fitur (*term*) dengan nilai QER tertinggi data *training* digunakan pada proses klasifikasi keempat. Jumlah fitur yang digunakan sebanyak 232 fitur.

Tabel 8. *Confusion Matrix* dengan 40% Fitur

Kelas Prediksi	Kelas Aktual	
	Negatif	Positif
Negatif	66	11
Positif	21	102

5. Penggunaan 20% Fitur

Sejumlah 20% fitur (*term*) yang memiliki nilai QER tertinggi pada data *training* digunakan pada proses klasifikasi terakhir. Jumlah fitur yang digunakan sebanyak 51 fitur.

Tabel 9. *Confusion Matrix* dengan 20% Fitur

Kelas Prediksi	Kelas Aktual	
	Negatif	Positif
Negatif	57	9
Positif	30	104

4.7.4. Perbandingan Hasil Kinerja Klasifikasi

Hasil kinerja klasifikasi *Multinomial Naïve Bayes* berupa nilai akurasi dan *kappa statistic* dengan berbagai jumlah fitur yang digunakan dapat dilihat dalam Tabel 10.

Tabel 10. Perbandingan Hasil Kinerja Klasifikasi

Jumlah Fitur	Akurasi	<i>Kappa Statistic</i>
100%	86,5%	72,43%
80%	89%	77,62%
60%	85%	69,24%
40%	84%	67,01%
20%	80,5%	59,2%

4.8. Text Visualisation

Word cloud pada Gambar 1 memperlihatkan bahwa kata yang sering digunakan dalam komentar diantaranya adalah “shopee”, “aplikasi”, “belanja”. Hal ini menunjukkan bahwa ulasan pengguna sesuai dengan topik aplikasi Shopee pada Harbolnas. *Word cloud* juga memperlihatkan kata “promo”, “sale”, “bagus”, “mudah” yang menunjukkan tanggapan positif pengguna mengenai berbagai promo yang diberikan oleh Shopee pada Harbolnas. Selain itu, terdapat pula kata “error” dan “lambat” yang menunjukkan bahwa beberapa pengguna juga mengeluhkan server Shopee yang sempat mengalami gangguan pada hari perayaan tersebut.

- McCallum, A. & Nigam, K. 1998. A Comparison of Event Models for Naïve Bayes Text Classification. 752. *Workshop on Learning for Text Categorization AAAI*: 1998.
- Mujilahwati, S. 2016. Pre-Processing Text Mining pada Data Twitter. *Seminar Nasional Teknologi Informasi dan Komunikasi 2016 (SENTIKA 2016)*: 18-19 Maret 2016.
- Parlar, T. & Özel, S.A. 2016. A New Feature Selection Method for Sentiment Analysis of Turkish Reviews. *International Symposium on INnovations in Intelligent SysTems and Applications (INISTA)*: Agustus 2016.
- Said, B. & Pranoto, Y.M. 2015. Klasifikasi Data SMS Center Bupati Pamekasan Menggunakan Naïve Bayes Dengan Mad Smoothing. *IdeaTech, STTS Surabaya*, Hal: 92-99.
- Tessem, B., Bjørnstad, S., Chen, W. & Nyre, L. 2015. Word Cloud Visualisation of Locative Information. *Journal of Location Based Services* Vol. 9, No. 4.
- Vierra & Garrett. 2015. Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine*, Vol. 37, No. 5, Hal: 360-363.
- Wahid, D. H. & Azhari, S.N. 2016. Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity. *Indonesian Journal of Computing and Cybernetics Systems*, Vol. 10, No. 2, Hal: 207-218.