

IMPLEMENTASI METODE *MULTINOMIAL NAÏVE BAYES CLASSIFIER* UNTUK ANALISIS SENTIMEN

Eva Liyan Woro Ningrum¹, Aris Puji Widodo²

^{1,2}Departemen Ilmu Komputer/Informatika, Fakultas Sains dan Matematika, Universitas Diponegoro
Jalan Prof. Soedarto, SH. Tembalang, Semarang, Indonesia, 50275
Email: ¹evaliyanworingrum@gmail.com

Abstract. GO-JEK is an information technology-based service consisting of 3 main services, namely GO-JEK, GO-LIFE, and GO-PAY. As time goes by, customer complaints arise about the lack of satisfaction with GO-JEK services. One measure of GO-JEK customer service satisfaction can be acknowledged by way of analyzing sentiments on the data of public opinion twett submitted on the twitter official account, @gojekindonesia. In this research, sentiment analysis was carried out using the Naïve Bayes Classifier learning algorithm. The number of data twitts used is 9987. The twit data is labeled according to the class which includes positive, negative and neutral classes. Then the next process is data preprocessing consisting of cleansing, tokenization, filtering, stemming, and stopword removal. The evaluation method using 10-fold cross validation with the results obtained is precision value of 80%, recall of 80%, f1-score of 80%, maximum accuracy of 82% and average accuracy of 79%.

Keywords: Analysis, Sentiment, Naïve Bayes Classifier, Twitter, GO-JEK

Abstrak. GO-JEK merupakan sebuah layanan berbasis teknologi informasi yang terdiri dari 3 layanan utama, yaitu GO-JEK, GO-LIFE, dan GO-PAY. Seiring dengan berjalannya waktu kemudian muncul keluhan pelanggan mengenai kurang puas dengan pelayanan GO-JEK. Salah satu mengukur kepuasan layanan pelanggan GO-JEK dapat dikukan dengan cara melakukan analisis sentimen terhadap data twett opini masyarakat yang disampaikan pada *official account twitter*, @gojekindonesia. Pada penelitian ini, dilakukan analisis sentimen menggunakan algoritma pembelajaran *Multinomial Naïve Bayes Classifier*. Jumlah data twitt yang digunakan sebanyak 9987. Data twit tersebut diberikan label sesuai dengan kelas yang meliputi kelas *positive*, *negative*, dan *neutral*. Kemudian proses berikutnya dilakukan *Preprocessing* data yang terdiri dari proses *cleansing*, tokenisasi, *filtering*, *stemming*, dan penghapusan *stopword*. Metode evaluasi menggunakan *10-fold cross validation* dengan hasil yang diperoleh adalah nilai *precision* sebesar 80%, *recall* sebesar 80%, *f1-score* sebesar 80%, akurasi maksimal sebesar 82% dan akurasi rata-rata sebesar 79%.

Kata kunci: Analisis, Sentimen, Naïve Bayes Classifier, Twitter, GO-JEK

I. PENDAHULUAN

Perkembangan teknologi informasi di Indonesia berkembang secara cepat, dalam hal ini juga berpengaruh pada ketersediaan sarana transportasi menggunakan ojek. Dengan memanfaatkan teknologi, maka sarana ojek di Indonesia saat ini sudah bertransformasi dari ojek konvensional menjadi ojek *online*. GO-JEK merupakan salah satu pelopor berdirinya ojek *online* di Indonesia. Layanan GO-JEK semakin berkembang, baik dari segi fasilitas yang ditawarkan

serta cakupan wilayah pelayanannya yang semakin luas. Pelayanan GO-JEK sering menjadi topik pembahasan dikalangan masyarakat baik secara langsung ataupun lewat internet. Beberapa masyarakat menyampaikan opini mereka tentang keunggulan GO-JEK, saran untuk beberapa pelayanan yang ada atau bahkan ungkapan kekecewaan mereka pada pelayanan GO-JEK yang mereka dapatkan. Salah satu wadah yang digunakan untuk menampung semua opini masyarakat mengenai GO-JEK adalah sosial media twitter. Setiap harinya banyak *tweet* opini yang dapat tertampung pada *official account* twitter GO-JEK. Banyaknya *tweet* opini dapat dijadikan peluang untuk menggali informasi mengenai penilaian dan evaluasi atas pelayanan GO-JEK yang sudah berjalan menggunakan analisis sentimen.

Analisis sentimen merupakan suatu bentuk teknologi yang salah satunya digunakan untuk melakukan analisis tingkat kepuasan layanan. Analisis sentimen adalah merupakan proses memahami, mengekstrak, dan mengolah data tekstual secara otomatis untuk mendapatkan informasi [1]. Dengan analisis sentimen, maka dari *tweet* opini masyarakat tentang GO-JEK dapat dilakukan klasifikasi ke dalam kelas sentimen menggunakan model klasifikator yang terbentuk dari proses pelatihan data latih yang tersedia.

Dalam melakukan analisis sentiment salah satunya harus dilakukan klasifikasi. Proses klasifikasi untuk menghasilkan sebuah model klasifikator banyak tersedia metode klasifikasi yang sering digunakan, akan tetapi metode klasifikasi yang digunakan pada penelitian adalah menggunakan *Multinomial Naïve Bayes Classifier*. Metode *Multinomial Naïve Bayes Classifier* adalah merupakan salah satu metode digunakan untuk melakukan klasifikasi dokumen berbasis teks, memiliki tingkat akurasi yang baik, dan waktu komputasinya lebih cepat [2]. Pada penelitian ini diharapkan dapat menghasilkan kualitas layanan GO-JEK dengan menggunakan pendekatan analisis sentimen, sehingga dapat memberikan deskripsi pelayanan yang sudah diberikan ke masyarakat, serta digunakan untuk perbaikan agar pelayanan GO-JEK dapat menjadi lebih baik.

II. ANALISIS SENTIMEN DAN NBC (REVIEW)

2.1 Analisis Sentimen

NLP adalah pembuatan program yang memiliki kemampuan untuk memahami bahasa manusia [3]. Prinsipnya bahasa alami adalah suatu bentuk representasi pesan yang ingin disampaikan, seringkali dalam bentuk suara dan tulisan. Sehingga dengan adanya NLP diharapkan pengguna dapat berkomunikasi dengan komputer menggunakan bahasa sehari-hari. Secara umum, jenis aplikasi yang dapat dibuat dalam bidang ilmu NLP terbagi dua yaitu *text-*

based application dan *dialogue-based application*. Beberapa alasan yang menyulitkan NLP adalah *ambiguity* dan jumlah kosa kata yang besar serta selalu berkembang [4].

Analisis sentimen adalah proses memahami, mengekstrak, dan mengolah data tekstual secara otomatis untuk mendapatkan informasi [1]. Selain itu analisis sentimen juga dapat diartikan sebagai bidang studi yang menganalisis opini, sentimen, evaluasi, penilaian, sikap dan emosi orang-orang terhadap entitas seperti produk, layanan, organisasi, individu, isu, peristiwa, topik, dan sifat-sifatnya [5]. Analisis sentimen dapat dianggap sebagai kombinasi antara *text mining* dengan NLP. Terdapat beberapa tingkat analisis yang dapat dilakukan untuk analisis sentiment, yaitu *document level*, *sentence level*, serta *entity and aspect level*.

2.2 Multinomial Naïve Bayes Classifier (NBC)

Bayesian classifier adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu *class*. *Bayesian classifier* terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar [6].

Dasar dari teorema naïve bayes pembentukan atribut vektor x^* yang dipakai dalam pemrograman adalah rumus Bayes: [7].

$$p(c|x^*) = \frac{p(x^*|c)p(c)}{p(x^*)} = \frac{p(x^*|c)p(c)}{\sum_c p(x^*|c)p(c)} \dots\dots\dots(1)$$

$$p(c) = \frac{\text{number of times class } c \text{ occurs}}{\text{total number of data points}} \dots\dots\dots(2)$$

Multinomial Naïve Bayes Classifier adalah algoritma pembelajaran yang dapat menangani banyak fitur dan dalam algoritma ini banyaknya kemunculan kata diperhitungkan, sehingga cocok untuk klasifikasi teks [8]. Metode *naïve bayes classifier* terdiri dari dua tahap dalam proses klasifikasi teks, tahap pelatihan dan tahap klasifikasi.

Dalam proses klasifikasi dokumen, terlebih dahulu melakukan penentuan kategori dari suatu dokumen berdasarkan kata-kata yang terkandung didalamnya dengan persamaan sebagai berikut [2]. Penentuan kategori (c) dari suatu dokumen (d).

$$c_{MAP} = \arg \max_{c \in \{c_1, c_s\}} P(c) = \arg \max_{c \in \{c_1, c_s\}} P(c) \prod_{k=1}^m P(t_k|c) \dots\dots\dots(3)$$

dimana t_k adalah kata dari dokumen d yang ingin diketahui kategorinya. nilai $p(t_k|c)$ diketahui dari data training yang dimiliki. Parameter $P(t_k|c)$ diestimasi dengan menghitung kejadian t_k pada semua dokumen training di c , menggunakan *Laplacean prior* [2].

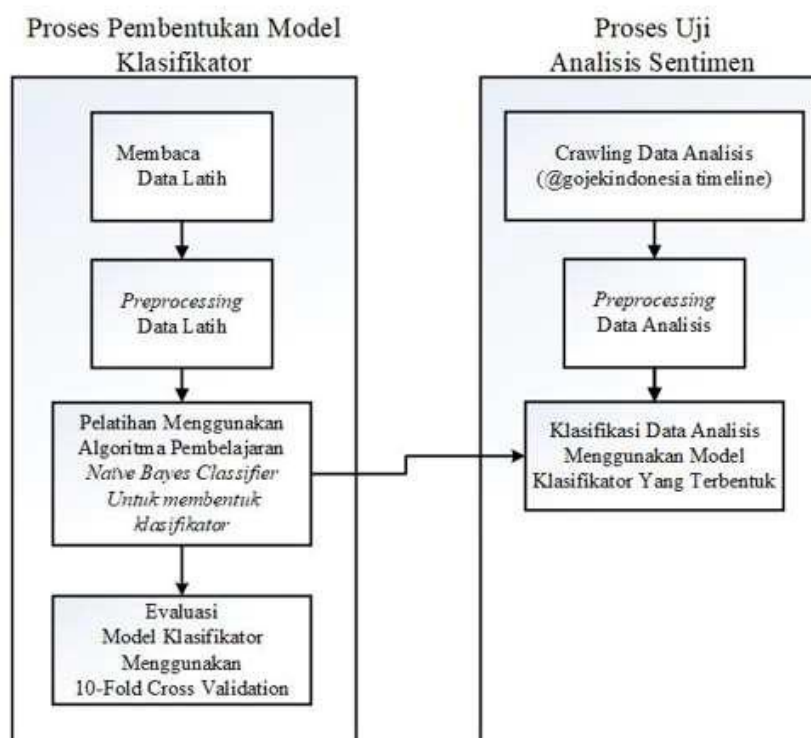
$$P(t_k|c) = \frac{1+N_k}{|V|+N'} \dots\dots\dots(4)$$

2.3 Data Preprocessing

Data *preprocessing* dilakukan untuk membersihkan dan mempersiapkan data sebelum dilakukan proses klasifikasi. Teks yang diambil dari web biasanya banyak mengandung *noise* dan teks yang tidak informatif, seperti tag-tag HTML, *script*, dan tanda baca yang tidak perlu. Tujuan dari data *preprocessing* adalah untuk membersihkan serta melakukan penyeragaman kata sehingga kata tersebut siap untuk diekstraksi ke tahap selanjutnya yaitu klasifikasi. *Preprocessing* dalam proses klasifikasi diperlukan untuk membangun *index* yang merepresentasikan isi atau topik dalam bentuk *term document* yang terkandung pada koleksi dokumen. Keluaran yang didapatkan melalui tahap *preprocessing* direpresentasikan sebagai bentuk *Bag of Words* (BoW) [9]. BoW adalah sebuah model yang merepresentasikan objek secara global misalnya kalimat teks atau dokumen sebagai bag (*multiset*) kata tanpa memperdulikan tata bahasa bahkan urutan kata untuk menjaga keanekaragamannya. Dengan kata lain, BoW merupakan kumpulan kata-kata unik dalam dokumen. Semua dokumen dapat direpresentasikan secara sederhana menggunakan BoW. Tahapan *preprocessing* yang dilakukan terdiri dari proses: *cleansing*, tokenisasi, *filtering*, *stemming*, dan penghapusan *stopword*.

III. METODE PENELITIAN

Tahapan penelitian ini dibagi menjadi 2 yaitu, proses pembentukan model klasifikator dan proses uji analisis sentiment, seperti yang diberikan pada Gambar 1. Proses pembentukan model klasifikator terdiri dari membaca 9987 data latih, *preprocessing* data latih, pelatihan dan evaluasi. Data latih yang digunakan berjumlah 9987 data latih (dengan pembagian sebanyak 3329 data sentimen *positive*, 3329 data sentimen *negative*, dan 3329 data sentimen *neutral*) berdasarkan data twett pada @gojekindonesia pada bulan Mei 2018 sampai bulan Juli 2018. Tahapan *preprocessing* data latih terdiri dari *cleansing*, tokenisasi, *filtering*, *stemming* menggunakan *library stemmer* Sastrawi, dan penghapusan *stopwords*.



Gambar 1. Tahapan penelitian

Tahapan proses pelatihan terdiri dari tiga langkah utama yaitu bentuk matriks *vocabulary*, hitung probabilitas setiap kelas sentimen $p(c)$, dan menentukan nilai frekuensi setiap *vocab* $\rightarrow P(t_k|c)$ pada setiap kelas. Untuk melakukan evaluasi kinerja klasifikator, aplikasi ASPG menggunakan *10-fold cross validation*. Tahapan proses uji analisis sentimen terdiri dari 3 tahap yaitu *crawling* data analisis, *preprocessing* data analisis, dan klasifikasi data analisis. Tahapan *preprocessing* data analisis pada penelitian ini terdiri dari *cleansing*, tokenisasi, *filtering*, *stemming*, dan penghapusan *stopword*. Tahapan proses klasifikasi data analisis terdiri dari 2 langkah proses utama yaitu menghitung nilai $C_{MAP} \rightarrow P(c) \prod_{k=1}^m P(t_k|c)$ berdasarkan model klasifikator yang sudah terbentuk dan hitung dan tentukan nilai C_{MAP} maksimal untuk menentukan *tweet* tersebut termasuk pada kelas sentimen tertentu.

IV. HASIL DAN PEMBAHASAN

Setelah dilakukan proses preprosesing dan dihasilkan sebuah analisis sentiment dari hasil *tweet*, maka tahapan berikut dilakukan proses pengujian kinerja. Pengujian kinerja dilakukan dengan dua eksperimen, dengan, skenario diberikan pada Tabel 1.

Tabel 1. Skenario pengujian kinerja sistem

No	Skenario	Tujuan
1	Melakukan eksperimen pengujian 9987 data latih dengan menerapkan 10-fold cross validation.	Eksperimen ini bertujuan untuk mengetahui besarnya akurasi dengan menerapkan 10-fold cross validation, dan bertujuan untuk mengetahui berapa nilai confusion matrix, precision, recall serta nilai akurasi maksimal yang tercapai.
2	Melakukan eksperimen analisis data dengan rentang waktu satu bulan dan data limit sebesar 500 data	Eksperimen ini bertujuan untuk mengetahui bagaimana tweet opini tentang GO-JEK dalam rentang waktu tertentu, apakah memiliki kecenderungan pada kelas sentimen negative, positive, atau neutral.

Tabel 2. Nilai hasil eksperimen skenario pertama

Hasil penerapan 10-fold cross validation			
Confusion Matrix	PPos	PNeg	PNeut
Aktual – Positive	579	52	69
Aktual – Negative	35	602	63
Aktual – Neutral	72	131	497
Precision			
Class Positif = $\frac{579}{579+35+72} = \frac{579}{686} = 0,84$			
Class Negatif = $\frac{602}{602+52+131} = \frac{602}{785} = 0,77$			
Class Netral = $\frac{497}{497+69+63} = \frac{497}{629} = 0,79$			
Mean Precision = $\frac{0,84+0,77+0,79}{3} = \frac{2,4}{3} = 0,80$			
Recall			
Class Positif = $\frac{579}{579+52+69} = \frac{579}{700} = 0,83$			
Class Negatif = $\frac{602}{35+602+63} = \frac{602}{700} = 0,86$			
Class Netral = $\frac{497}{72+131+497} = \frac{497}{700} = 0,71$			
Mean Recall = $\frac{0,83+0,86+0,71}{3} = \frac{2,4}{3} = 0,80$			
F1-score			
Class Positif = $\frac{2 \times 0,84 \times 0,83}{0,84 + 0,83} = \frac{1,3944}{1,67} = 0,84$			
Class Negatif = $\frac{2 \times 0,77 \times 0,86}{0,77 + 0,86} = \frac{1,3244}{1,63} = 0,81$			
Class Netral = $\frac{2 \times 0,79 \times 0,71}{0,79 + 0,71} = \frac{1,1218}{1,5} = 0,75$			
Mean f1-score = $\frac{0,84+0,81+0,75}{3} = \frac{2,4}{3} = 0,80$			
Akurasi setiap iterasi pada 10-fold cross validation			
Iterasi 1 = 0,82857			
Iterasi 2 = 0,75714			
Iterasi 3 = 0,82380			
Iterasi 4 = 0,82380			
Iterasi 5 = 0,76666			
Iterasi 6 = 0,79523			
Iterasi 7 = 0,81904			
Iterasi 8 = 0,78571			
Iterasi 9 = 0,79407			
Iterasi 10 = 0,8			
Akurasi rata-rata			
0,79904			
Akurasi maksimal			
0,82857			

Untuk hasil eksperimen skenario pertama dari penerapan *10-fold cross validation* memiliki nilai tingkat akurasi sebesar 79%, *precision* 80%, *recall* 80%, dan *f1-score* 80%, seperti yang diberikan pada Tabel 2.

Berdasarkan hasil eksperimen pertama dapat dilakukan analisis sebagai berikut:

- a. Memiliki tingkat keakuratan 79% dari keseluruhan data latih yang dimiliki dalam keberhasilannya melakukan analisis sentimen pelayanan gojek secara tepat. Mampu melakukan analisis sentimen pelayanan gojek dan memprediksinya secara tepat pada tiga kelas sentimen yang ada yaitu *positive*, *negative*, dan *neutral* dibuktikan dengan nilai akurasi yang dicapai sebesar 79%.
- b. Memiliki tingkat kebenaran 80% untuk keberhasilannya dalam melakukan analisis sentimen pada data latih secara benar dan tepat dibandingkan dengan nilai analisis sentimen aktual atau sebenarnya yang dilakukan secara manual, hal ini dibuktikan dengan nilai *precision* yang dapat dicapai sebesar 80%.
- c. Memiliki tingkat keberhasilan 80% untuk melakukan analisis sentimen pada data analisis secara benar dan tepat, hal ini dibuktikan dengan nilai *recall* yang dapat dicapai sebesar 80%.
- d. Memiliki nilai rata-rata keakuratan 80% untuk melakukan analisis sentimen pada data latih secara tepat dibandingkan dengan nilai aktualnya serta dapat melakukan analisis sentimen secara benar dan tepat, hal ini dibuktikan dengan *f1-score* yang dapat dicapai sebesar 80%.

Skenario hasil eksperimen kedua menunjukkan hasil analisis sentimen pelayanan GO-JEK pada twitter dengan data analisis menghasilkan kecenderungan *tweet* opini masyarakat mengenai pelayanan GO-JEK unggul pada kelas sentimen *neutral* dengan nilai presentase 55,8%, sedangkan kelas sentimen *positive* 10,8%, dan kelas sentimen *negative* 33,5% dari total 800 data analisis yang diproses. Setelah dilakukan analisa hasil menunjukkan bahwa rata-rata data analisis yang didapat adalah *tweet* opini seputar pertanyaan mengenai pelayanan GO-JEK dan *tweet* opini seputar jawaban penjelasan dari pihak *customer service* GO-JEK. Untuk mengukur kinerja lebih lanjut, dalam analisis hasil ini dilakukan perbandingan terhadap analisis sentimen yang dihasilkan pada penelitian ini dengan analisis sentimen aktual yang dilakukan secara manual dan menghasilkan nilai akurasi sebesar 84%, *precision* 70%, *recall* 88%, dan *f1-score* 74%, seperti yang diberikan pada Tabel 3.

Tabel 3. Nilai hasil eksperimen skenario pertama

Analisis hasil eksperimen kedua			
Confusion Matrix	PPos	PNeg	PNeut
Aktual – Positive	33	2	0
Aktual – Negative	13	200	7
Aktual – Neutral	41	64	440
Precision			
Class Positif	$= \frac{33}{33+13+41} = \frac{33}{87} = 0,37$		
Class Negatif	$= \frac{200}{200+2+64} = \frac{200}{266} = 0,75$		
Class Netral	$= \frac{440}{440+7+0} = \frac{440}{447} = 0,98$		
Mean Precision	$= \frac{0,37+0,75+0,98}{3} = \frac{2,1}{3} = 0,7$		
Recall			
Class Positif	$= \frac{33}{33+2+0} = \frac{33}{35} = 0,94$		
Class Negatif	$= \frac{200}{13+200+7} = \frac{200}{220} = 0,90$		
Class Netral	$= \frac{440}{41+64+440} = \frac{440}{545} = 0,80$		
Mean Recall	$= \frac{0,94+0,90+0,80}{3} = \frac{2,64}{3} = 0,88$		
F1-score			
Class Positif	$= \frac{2 \times 0,37 \times 0,94}{0,37 + 0,94} = \frac{0,6956}{1,31} = 0,53$		
Class Negatif	$= \frac{2 \times 0,75 \times 0,90}{0,75 + 0,90} = \frac{1,35}{1,65} = 0,8$		
Class Netral	$= \frac{2 \times 0,98 \times 0,80}{0,98 + 0,80} = \frac{1,568}{1,78} = 0,88$		
Mean f1-score	$= \frac{0,53+0,8+0,88}{3} = \frac{2,22}{3} = 0,74$		
Akurasi			
Akurasi	$= \frac{33+200+440}{800} = \frac{673}{800} = 0,84$		

Berdasarkan hasil eksperimen kedua dapat dilakukan analisis sebagai berikut:

- a. Memiliki tingkat keakuratan 84% dalam keberhasilannya melakukan analisis sentimen pelayanan gojek secara tepat. Mampu melakukan analisis sentiment dan berhasil memprediksi data analisis pelayanan GO-JEK secara tepat masuk kedalam kelas sentimen yang sesuai, hal ini diperlihatkan dengan nilai akurasi yang dicapai pada eksperimen kedua sebesar 84%.
- b. Memiliki tingkat kebenaran 70% untuk keberhasilannya melakukan analisis sentimen dan mengklasifikasikan data secara benar dan tepat dibandingkan dengan hasil analisis sentimen aktual atau sebenarnya yang dilakukan secara manual, hal ini dibuktikan dengan nilai *precision* yang dapat dicapai pada eksperimen kedua sebesar 70%.
- c. Memiliki tingkat keberhasilan 80% untuk melakukan analisis sentimen pada data analisis mengenai pelayanan GO-JEK secara benar dan tepat. Penelitian ini berhasil melakukan analisis sentimen dan mengklasifikasikan data analisis pelayanan GO-JEK pada kelas sentimen yang benar dan tepat hal ini dibuktikan dengan nilai *recall* yang dapat dicapai sebesar 80%.

d. Memiliki nilai rata-rata keakuratan 80% untuk melakukan analisis sentimen data analisis pelayanan GO-JEK secara tepat dibandingkan dengan nilai aktualnya serta dapat melakukan analisis sentimen dan mengklasifikasikan data analisis sesuai kelas sentimennya secara benar dan tepat, hal ini dibuktikan dengan *f1-score* yang dapat dicapai sebesar 80%.

V. KESIMPULAN

Berdasarkan hasil eksperimen yang telah dilakukan diatas, maka dapat diberikan kesimpulan bahwa hasil kinerja model klasifikator yang terbentuk dari proses pelatihan 9987 data latih diperoleh nilai akurasi rata-rata 79%, nilai akurasi maksimal 82%, nilai *precision* 80%, nilai *recall* 80% dan *f1-score* 80%. Nilai akurasi yang dicapai dalam penelitian ini terpengaruhi oleh faktor jumlah data latih yang digunakan, keberagaman jenis data latih, proses *preprocessing* yang digunakan, daftar *stopword*, daftar *anomaliwords* yang digunakan, dan metode pengujian klasifikator yang digunakan. Analisis sentimen diperoleh kecenderungan *tweet* opini mengenai pelayanan GO-JEK unggul pada kelas sentiment *neutral*. Presentase hasil untuk kelas *neutral* mencapai 55,8%, sedangkan sentimen *positive* 10,8%, dan sentimen *negative* 33,5%.

DAFTAR PUSTAKA

- [1] Pang, Bo. & Lee, Lillian. 2008. *Opinion Mining and Sentiment Analysis*. Foundation and Trends in Information Retrieval, vol. 2, no 1-2, pp. 1-135..
- [2] Berry, M. W., & Kogan, J. 2010. *Text Mining Application and Theory*. United Kingdom: John Wiley and Sons.
- [3] Sutojo, T. Mulyanto Edy dan Suhartono Vincent. 2011. *KECERDASAN BUATAN*. Yogyakarta: Andi Offset..
- [4] Lisangan, Erick Alfons. 2013. *Natural Language Processing Dalam Memperoleh Informasi Akademik Mahasiswa Universitas Atma Jaya Makassar*. Jurnal Tematika Vol.1, No.1, ISSN: 23033878..
- [5] Liu, B. 2012. *Sentiment Analysis and Opinion Mining*. s.l.:Morgan & Claypool Publishers.
- [6] Luthfi, E. T., & Kusriani. 2009. *Algoritma Data Mining*. Yogyakarta: ANDI..
- [7] Barber, D.2012. *Bayesian Reasoning and Machine Learning*. United Kingdom: Cambridge University Press.
- [8] Manning, C. D., Raghavan, P., & Schütze, H. 2009. *Introduction to Information Retrieval*. Cambridge: Cambridge University Press..

- [9] Verbiest, N., Vermeulen, K., dan Teredesai, A. 2015. *Evaluation of Classification Methods*, In *Data Classification Algorithms and Applications*, by Aggarwal, C.C., 633-652. New York: CRC Press