

Pre-Processing Kalsifikasi Karakter Baik Buruk Seseorang Menggunakan Data Twitter

Evi Pertiwi Munthe¹, Rima Dias Ramadhani², Muhammad Zidny Naf'an³, Rifki Adhitama⁴

^{1,2,3} Program Studi Informatika, ⁴ Program Rekayasa Perangkat Lunak,

^{1,2,3,4} Fakultas Teknologi Industri dan Informatika,

^{1,2,3,4} Institut Teknologi Telkom Purwokerto

^{1,2,3,4} Jl. D.I Panjaitan No. 128, Purwokerto Kidul, Purwokerto Selatan, Purwokerto, 53147, Jawa Tengah, Indonesia

Email : ¹14102061@ittelkom-pwt.ac.id, ²rima@ittelkom-pwt.ac.id, ³zidny@ittelkom-pwt.ac.id, ⁴rifki@ittelkom-pwt.ac.id

Abstrak – Jumlah pengguna Twitter di Indonesia mencapai 332 juta pengguna per bulannya pada tahun 2015. Twitter merupakan sebuah situs web yang menawarkan jejaringan sosial dalam bentuk *microblog*. Dari *microblog* pada twitter dapat diketahui karakter seseorang. Salah satu penelitian untuk mengetahui karakter dari seseorang berdasarkan data twitter di antaranya adalah “*You are what you tweet: Personality expression and perception on twitter*”. Pada penelitian ini membahas tentang hubungan antara karakter seseorang dan media social berdasarkan *microblog* pada twitter dengan analisis linguistik. Pada penelitian menggunakan teknik pencocokan kosa kata yang digunakan oleh pengguna twitter dengan kamus data yang tergolong baik atau buruk secara manual dengan menggunakan teknik analisis linguistik. Hasil dari penelitian tersebut, didapatkan klasifikasi baik buruk yang memiliki akurasi yang baik. Namun dalam studi tersebut belum ada algoritma *pre-processing* yang dapat digunakan untuk menentukan karakter dari seseorang untuk mempercepat klasifikasi dengan hasil akurasi yang baik. Teknik *pre-processing* digunakan untuk menghilangkan noise data tweet yang sudah diperoleh. Namun dalam penelitian ini hasil dari *pre-processing* kurang akurat dikarenakan library dan kamus yang digunakan dalam mengelola data kurang lengkap. Penggunaan fitur *tf idf* dapat memberikan informasi seberapa banyak kata tersebut muncul dalam dokumen yang sudah memiliki klasifikasi baik dan buruk netral pada masing-masing kriteria.

Kata kunci – Twitter, *tweet*, *pre-processing*, klasifikasi

Abstract— *The number of Twitter users in Indonesia reached 332 million users per month in the year 2015. Twitter is a website that offers a social network in the form of microblog. From the microblog on twitter can be known to someone's character. One of the research to find out the character of a person based on twitter data of which is "You are what You tweet: Personality expression and perception on twitter". This research discusses the relationship between the character of a person and social media based on microblog on twitter with a linguistic analysis. On the research of using the technique of matching the vocabulary used by the twitter user with a data dictionary belongs to good or bad manually by using linguistic analysis techniques. The results of such research, obtained good bad classification which has a good accuracy. But in the study there is no pre-processing algorithms that can be used to determine the character of a person is to speed up the classification with a good accuracy results. Pre-processing techniques used to remove noise that tweet data already acquired. However, in this study the results of pre-processing less accurate due to the library and dictionary which is used to manage the data less complete. Feature use tf idf can provide information how much those words appear in the document already has a good and bad neutral classification on each criterion.*

Keywords— *Twitter, Tweet, pre-processing, classification*

I. PENDAHULUAN

Perkembangan teknologi yang semakin pesat mempengaruhi pertumbuhan jumlah data. Sekarang ini, data merupakan sesuatu yang penting dalam berbagai bidang, mulai dari bidang akademik, bisnis

dan media [1]. Hal ini disebabkan adanya kebutuhan organisasi yang tinggi untuk mendapatkan informasi yang berguna dalam pengambilan keputusan [1] [2].

Pengguna twitter di Indonesia mencapai 332 juta pengguna per-bulannya. Pada pertengahan tahun 2015

pengguna twitter mencapai 50 juta. 77 persen pengguna twitter di Indonesia aktif setiap harinya. Dari 77 persen, 54 persen diantaranya melakukan 2 *tweet* setiap harinya. Konten video juga mulai banyak digunakan di Indonesia yaitu 44 persen pengguna yang menggunakan video saat melakukan *tweet* dan 43 persen diantaranya melakukan *retweet* dari video yang ada [3].

Twitter merupakan sebuah situs *web* yang menawarkan jejaring sosial dalam bentuk *microblog*, dimana pengguna dapat mengirimkan dan membaca pesan yang disebut dengan *tweet*. Kapasitas *tweet* adalah sebanyak 280 karakter, hal ini menyebabkan pengguna harus mengekspresikan buah pikiran mereka dalam suatu kalimat kunci [4]. Pengguna twitter menggunakan *microblog* untuk menggambarkan rutinitas sehari-hari mereka, melakukan percakapan, melaporkan berita dan berbagi informasi. Dari *microblog* pada twitter dapat diketahui karakter seseorang. Terdapat berbagai penelitian untuk mengetahui karakter dari seseorang berdasarkan data twitter di antaranya adalah “*You are what you tweet: Personality expression and perception on twitter*”[5]. Pada penelitian ini membahas tentang hubungan antara karakter seseorang dan media *social* berdasarkan *microblog* pada twitter dengan analisis linguistik. Pada penelitian [5] menggunakan teknik pencocokan kosa kata yang digunakan oleh pengguna twitter dengan kamus data yang tergolong baik atau buruk secara manual dengan menggunakan teknik analisis linguistik. Hasil dari penelitian tersebut, didapatkan klasifikasi baik buruk yang memiliki akurasi yang baik. Namun dalam studi tersebut belum ada algoritma *pre-processing* yang dapat digunakan untuk menentukan karakter dari seseorang. Dalam metode *pre-processing* terdapat berbagai metode yang telah diimplementasikan untuk mengklasifikasi dengan tingkat akurasi yang didapatkan tinggi [6-12].

II. METODE PENELITIAN

A. Pengumpulan Data

Adapun tahap-tahap yang harus dilakukan pada pengumpulan data adalah mengambil data dari sebuah penelitian yang berjudul tentang “Prediksi Profesi Berdasarkan Model Bahasa Pada *Tweets*”, tahun 2016 oleh Hapnes Toba. Pada penelitian ini data terdapat 4 kelas yang sudah diklasifikasikan yaitu kelas artis, kelas pelajar, kelas politik dan kelas musisi. Namun pada penelitian ini klasifikasi dilakukan hanya menggunakan 2 kelas yaitu baik dan buruk, sehingga untuk mendapatkan klasifikasi data tersebut dilakukan validasi data.

B. Validasi Data

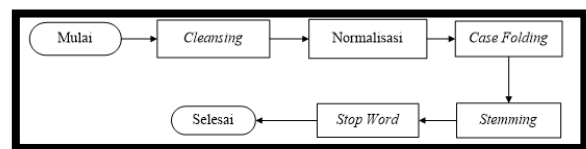
Pada tahap ini dilakukan pemilihan *tweet* yang akan digunakan. Adapun *tweet* yang diambil hanya *tweet* yang menggunakan bahasa Indonesia saja. Selanjutnya melakukan pencarian *tweet* yang memiliki kesesuaian dengan karakteristik data dalam

menentukan karakter baik atau buruk. Setelah itu dilakukan klasifikasi manual dengan cara membuat dalam bentuk kuisioner. Hasil dari rekapan kuisioner dilakukan perhitungan untuk menentukan data valid. Melakukan perhitungan keakuratan data dengan menggunakan software SPSS.

C. Pre-Processing

Pada tahap ini dilakukan kajian pustaka untuk mempermudah *pre-processing*. Pustaka yang digunakan pada proses ini adalah algoritma dari sastrawi master dengan menggunakan bahasa pemrograman PHP. *Pre-processing* sendiri memiliki beberapa tahap yang harus dilakukan agar memperoleh data yang sesuai.

Berikut ini tahap-tahap *pre-processing* yang dilakukan.



Gambar 1. Flowchart Pre-Processing

a) Cleansing

Cleansing merupakan proses yang dilakukan untuk menghilangkan semua tanda baca yaitu tanda titik (.), tanda koma (,), dan seru (!), tanda tanya (?) dan lain-lain. *Cleansing* juga dapat digunakan untuk menghilangkan kata-kata yang tidak diperlukan pada klasifikasi seperti *hashtag* (#), URL dan *username* (@), namun pada proses ini dilakukan pada MS.Excel.

b) Normalisasi

Mengubah kata yang tidak resmi atau gaul dan singkatan ke dalam bentuk kata yang resmi menurut Kamus Besar Bahasa Indonesia (KBBI). Kamus yang dibuat pada penelitian ini adalah kamus manual dengan menggunakan format JSON. Program akan membaca kalimat yang di input lalu setiap kata pada kalimat akan dibaca program lalu disesuaikan dengan kamus manual yang telah dibuat.

c) Stemming

Algoritma yang digunakan pada tahap stemming ini adalah algoritma stemming Nazief-Adriani, yang kemudian disempurnakan oleh algoritma CS (*Confix Stripping*), kemudian disempurnakan oleh ECS (*Enhanced Confix Stripping*) dan kemudian disempurnakan lagi oleh Modified ECS. Algoritma ini bekerja dengan cara setiap kalimat akan diproses dengan mengubah kata berimbuhan menjadi kata dasar dengan cara menghilangkan awalan maupun akhiran dan bisa juga keduanya.

d) Stop Word

Dengan menggunakan fitur ini kalimat sebelum diklasifikasi terlebih dahulu dengan menghilangkan kalimat yang tidak berhubungan dengan klasifikasi

sehingga hasilnya kalimat akan berkurang namun tidak mengurangi makna dari isi kalimat.

e) *Tokenisasi*

Kalimat akan dilakukan pemecahan menjadi bentuk token-token menggunakan pembatas atau sering disebut delimiter spasi. Untuk nentukan nilai bobot dari setiap token-token dilakukan perhitungan TF-IDF. TF atau term frequency adalah jumlah kemunculan teks atau term dalam dokumen [16]. Berikut ini persamaan TF.

$$W(d,t) = TF(d,t) \dots\dots\dots(1)$$

IDF merupakan jumlah dokumen yang terdapat *term* [20]. Berikut ini persamaan IDF.

$$Idf(t) = \log \frac{N}{df(t)} \dots\dots\dots(2)$$

Sehingga dari persamaan diatas maka TF-IDF merupakan kombinasi dari metode TF dan IDF. Berikut ini persamaan TF-IDF.

$$TF-IDF(d,t) = TF(d,t)*IDF(t) \dots\dots\dots(3)$$

III. HASIL PENELITIAN

Untuk mendapatkan hasil penelitian ini dapat diperoleh dengan melakuakan pengujian. Berikut ini hasil dari pengujian dari masing-masing tahap penelitian.

A. Validasi Data

Dalam pengelolaan data ada beberapa pengujian yang dilakukan yaitu terlebih dahulu dilakukan klasifikasi *tweet* ke dalam baik dan buruk dilakukan secara manual untuk mendapatkan data latih. *Tweet* yang diperoleh yaitu 550 *tweet*. Adapun jumlah responden yang dilakukan adalah sebanyak 52 responden. Dengan persentase umur 16-20 tahun sebanyak 13 %, umur 21-25 tahun sebanyak 75 % dan umur 26-30 tahun sebanyak 12 %. Adapun persentase responden yang menggunakan twitter sebanyak 88 % dan yang tidak menggunakan twitter sebanyak 12 %. Persentase seberapa sering menggunakan twitter yaitu sebanyak 4 % sangat sering menggunakan twitter, 2 % sering menggunakan, 15 % kadang-kadang menggunakan, 63 % menggunakan jarang dan 15 % tidak pernah menggunakan twitter. Setelah dilakukan penjumlahan dari hasil responden maka diketahui 432 *tweet* masuk kedalam klasifikasi *tweet* baik, 114 *tweet* masuk kedalam klasifikasi *tweet* buruk dan 4 *tweet* netral. Karena dalam penelitian ini yang digunakan hanya klasifikasi *tweet* baik dan buruk maka hasil responden yang netral tidak dimasukkan dalam klasifikasi ini. Sehingga *tweet* yang digunakan adalah 432 *tweet* baik dan 114 *tweet* buruk.

B. Perhitungan Statistik Data

Sebelum melakukan *pre-processing*, terlebih dahulu harus diketahui kelayakan data untuk diolah. Hasil dari pengujian menggunakan software SPSS, diketahui bahwa data sudah *valid* dan layak untuk diolah. Hasil dari SPSS menunjukkan nilai .944.

Menurut penelitian [14] nilai .944 termasuk dalam kategori *excellent* karena hasil yang dicapai ≥ 0.9 .

Tabel 1. Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.944	.946	52

C. Pelabelan Data

Berdasarkan hasil data responden maka dilakukan pemberian bobot dengan menggunakan Ms. Excel. Pemberian bobot dilakukan berdasarkan *tweet* yang akan dinilai yaitu *tweet* mengenai kata-kata baik akan diberi nilai 1 dan kata-kata buruk nilai -1 dikarenakan tidak adanya penentuan *tweet* yang mengandung kata netral (kata-kata baik dan kata-kata tidak baik) yang menjadi dasar penentuan karakter baik dan buruk seseorang pada penelitian ini. Berikut ini tampilan bobot *tweet*.

Tabel 2. Pelabelan Data

ID	<i>Tweet</i>	Resp 1	Resp 2
5	Selamat hari senin, selamat beraktifitas kerja dan kuliah bagi para dosen dan mahasiswa daerah maranatha [pic]	1	1
269	Sebelum tidur jangan lupa berdoa dulu. Doanya sesuai keyakinan masing2 aja, gak usah diseragamin.	1	1
7	Gw suka sombong sih, sekarang suka ada aja org yg sombong ckck	-1	-1
96	Yang timnya menang, santai aja gak usah ngehina.	-1	-1

Dengan dilakukannya pembobotan secara menyeluruh dengan jumlah 52 responden maka selanjutnya dilakukan perhitungan nilai bobot yang akan menentukan label *tweet* yaitu *tweet* baik dan *tweet* yang buruk. Setelah ditemukan label *tweet* baik dan *tweet* buruk dilakukan pemisahan kedua label *tweet* tersebut dengan 2 folder yaitu folder *tweet* baik dan *tweet* buruk. Dari hasil pelabelan ditemukan jumlah *tweet* baik sebanyak 432 *tweet*, *tweet* buruk sebanyak 114 *tweet* dan *tweet* netral sebanyak 4 *tweet*. Pada penelitian ini yang diambil hanya *tweet* baik dan *tweet* buruk saja.

D. *Pre-processing*

Setelah dilakukan pelabelan secara manual, selanjutnya dilakukan *pre-processing*. Pada *pre-processing* ini ada beberapa tahap yang digabungkan dalam output sehingga ada dua hasil output yang dihasilkan dalam *pre-processing* ini yaitu yang

pertama tahap *cleansing* yaitu untuk menghilangkan tanda baca maupun angka dan normalisasi yaitu mengubah kata gaul menjadi kata baku berdasarkan KBBI. Pada tahap normalisasi ini dilakukan pembuatan kamus secara manual lalu di simpan dalam format JSON. Berikut ini hasil dari *pre-processing* pada tahap pertama.

Tabel 3. Hasil *Pre-Processing* Tahap 1

ID	<i>Tweet</i> Sebelum <i>Pre-processing</i>	Kamus Manual	<i>Pre-processing</i> Tahap 1
5	Selamat hari senin, selamat beraktifitas kerja dan kuliah bagi para dosen dan mahasiswa daerah maranatha [pic]		Selamat hari senin selamat aktifitas kerja dan kuliah bagi para dosen dan mahasiswa daerah maranatha
269	Sebelum tidur jangan lupa berdoa dulu. Doanya sesuai keyakinan masing2 aja, gak usah diseragamin.	aja = saja, gak= tidak,	Sebelum tidur jangan lupa doa dulu Doanya sesuai yakin masing saja tidak usah diseragamin
7	Gw suka sombong sih, sekarang suka ada aja org yg sombong ckck	Gw = saya, aja = saja, org = orang sih = " ", yg = yang, ckck = " "	Saya suka sombong sekarang suka ada aja orang yang sombong
96	Yang timnya menang, santai aja gak usah ngehina.	aja = saja, gak= tidak,	Yang tim menang santai saja tidak usah hina

Pada kedua yaitu *case folding* yaitu mengubah semua huruf pada *tweet* dari huruf kapital menjadi huruf kecil, *stop word* yang digunakan untuk menghapus kalimat yang tidak berhubungan dengan klasifikasi dan *stemming* yang digunakan untuk setiap kalimat akan diambil kata dasar saja. Berikut ini hasil *pre-processing* tahap yang kedua.

Tabel 4. Hasil *Pre-Processing* Tahap 2

ID	<i>Tweet</i> Sebelum <i>Pre-processing</i>	<i>Pre-processing</i> Tahap 2
5	Selamat hari senin selamat aktifitas kerja dan kuliah bagi para dosen dan mahasiswa daerah maranatha	selamat senin selamat aktivitas kuliah dosen mahasiswa daerah Maranatha
269	Sebelum tidur jangan lupa doa dulu Doanya sesuai yakin	belum tidur lupa doa doa sesuai yakin masing diseragamin

	masing saja tidak usah diseragamin	
7	Saya suka sombong sekarang suka ada aja orang yang sombong	saya suka sombong suka sombong
96	Yang tim menang santai saja tidak usah ngehina	yang tim menang santai ngehina

E. Tokenisasi dan TF-IDF (*Term Frequency-Inverse Document Frequency*)

Tokenisasi merupakan proses untuk membagi atau pemecahan teks dapat berupa kalimat atau *file* menjadi token-token.

Tabel 5. Hasil Tokenisasi

ID	<i>Tweet</i> Sebelum Tokenisasi	Sesudah Tokenisasi
5	Selamat hari senin selamat aktifitas kerja dan kuliah bagi para dosen dan mahasiswa daerah maranatha	"selamat", "senin", "selamat", "aktivitas", "kuliah", "dosen", "mahasiswa", "daerah", "Maranatha"
269	Sebelum tidur jangan lupa doa dulu Doanya sesuai yakin masing saja tidak usah diseragamin	"belum ", "tidur", "lupa", "doa", "doa", "sesuai", "yakin", "masing", "diseragamin"
7	Saya suka sombong sekarang suka ada aja orang yang sombong	"saya", "suka", "sombong", "suka", "sombong"
96	Yang tim menang santai saja tidak usah hina	"yang", "tim", "menang", "santai", "ngehina"

Metode TF-IDF merupakan metode yang digunakan untuk pembobotan term yang banyak yang digunakan untuk perbandingan terhadap pembobotan baru.

Tabel 6. Hasil TF-IDF

Term	TF				IDF		TF*IDF
	A	B	C	D	Df	N/df	
Selamat	2	0	0	0	1	0.60206	1.20412
Senin	1	0	0	0	1	0.60206	0.60206
Aktivitas	1	0	0	0	1	0.60206	0.60206
Kuliah	1	0	0	0	1	0.60206	0.60206
Dosen	1	0	0	0	1	0.60206	0.60206
Mahasiswa	1	0	0	0	1	0.60206	0.60206
Daerah	1	0	0	0	1	0.60206	0.60206
Maranatha	1	0	0	0	1	0.60206	0.60206
Belum	0	1	0	0	1	0.60206	0.60206

Tidur	0	1	0	0	1	0.60206	0.60206
Lupa	0	1	0	0	1	0.60206	0.60206
Doa	0	2	0	0	1	0.60206	1.20412
Sesuai	0	1	0	0	1	0.60206	0.60206
Yakin	0	1	0	0	1	0.60206	0.60206
Masing	0	1	0	0	1	0.60206	0.60206
Diseragami n	0	1	0	0	1	0.60206	0.60206
Saya	0	0	1	0	1	0.60206	0.60206
Suka	0	0	2	0	1	0.60206	1.20412
Sombong	0	0	2	0	1	0.60206	1.20412
Yang	0	0	0	1	1	0.60206	0.60206
Tim	0	0	0	1	1	0.60206	0.60206
Menang	0	0	0	1	1	0.60206	0.60206
Santai	0	0	0	1	1	0.60206	0.60206
Ngehina	0	0	0	1	1	0.60206	0.60206

IV. PEMBAHASAN

Pada penelitian ini ada beberapa tahapan yang dilakukan sebelum melakukan klasifikasi agar hasil klasifikasi lebih akurat yaitu tahap *pre-processing* pada data yang akan diolah. Pada penelitian ini data yang diolah yaitu berupa teks yang diambil dari tweet yang ada pada Twitter. Adapun tahap *pre-processing* yang dilakukan yaitu menghilangkan tanda baca atau angka, mengubah huruf kapital menjadi huruf kecil, mengubah kata gaul menjadi kata yang berdasarkan Kamus Besar Bahasa Indonesia (KBBI) secara manual, mengubah kata yang berimbuhan menjadi kata dasar dan menghilangkan kata-kata yang tidak memiliki arti dan tidak mengubah makna dari kalimat tersebut. Pada tahap ini, hasil data yang diperoleh masih kurang akurat dikarenakan pada tahap mengubah kata berimbuhan menjadi kata dasar dan menghilangkan kata yang tidak memiliki arti, program yang dijalankan belum secara menyeluruh dapat membaca kalimat yang diinputkan dan *library* yang digunakan pada *stemming* dan kamus manual masih kurang lengkap. Dalam tahap *pre-processing*, tidak membutuhkan waktu yang lama dalam pengeksekusian program.

Dari tabel 6 diketahui bahwa semakin banyak jumlah term pada dokumen maka nilai dari TF-IDF akan semakin besar. Sedangkan semakin banyak atau besar nilai DF maka nilai dari IDF semakin kecil. Sebagai buktinya pada kata "selamat" memiliki *term* kemunculan sebanyak 2 sehingga nilai dari TF-IDF lebih besar dari nilai TF-IDF kata lainnya.

V. PENUTUP

Pre-processing dalam klasifikasi sangat berpengaruh. Semakin akurat *pre-processing* maka dalam melakukan klasifikasi pada data juga semakin akurat. Keakuratan *pre-processing* ditentukan dengan *library* dan kamus yang digunakan. Semakin lengkap *library* dan kamus yang digunakan maka tingkat akurasi *pre-processing* semakin meningkat. Dalam penggunaan fitur *tf idf* dapat memberikan informasi seberapa banyak kata tersebut muncul dalam dokumen yang sudah memiliki klasifikasi baik dan buruk netral pada masing-masing kriteria.

DAFTAR PUSTAKA

- [6] E. Prasetyo, DATA MINING-Mengolah Data menjadi Informasi Menggunakan Matlab, Yogyakarta: Andi, 2014.
- [7] E. Prasetyo, DATA MINING-Konsep dan Aplikasi Menggunakan MATLAB, Yogyakarta: ANDI, 2012.
- [8] A. Maulana, "www.cnnindonesia.com," CNN Indonesia, 23 Maret 2016. [Online]. Available: <http://www.cnnindonesia.com/teknologi/20160322085045-185-118939/twitter-rahasiakan-jumlah-pengguna-di-indonesia/>. [Accessed 27 April 2017].
- [9] N.Y. A. Faradhillah, R. P. Kusumawardani, dan I. Hafidz, "Eksperimen Sistem Klasifikasi Analisa Sentimen Twitter pada Akun Resmi Pemerintah Kota Surabaya Berbasis Pembelajaran Mesin (Experiments on Sentiment Classification System for Tweets of the Official Account of the City Government of Surabaya based on Mach)," Pros. Semin. Nas. Sist. Inf. Indones. 2016, pp. 15–24, 2016.
- [10] L. Qiu, H. Lin, J. Ramsay dan F. Yang, "You are what you tweet: Personality expression and perception on Twitter," Journal of Research in Personality, vol. 46, no. 6, pp. 710-718, 2012.
- [11] S. Urolagin, "Text Mining of Tweet for Sentiment Classification and Association with Stock Prices," 2017 Int. Conf. Comput. Appl. ICCA 2017, pp. 384–388, 2017.
- [12] I. B. N. S. Hardaya, A. Dhini, dan I. Surjandari, "Application of text mining for classification of community complaints and proposals," 2017 3rd Int. Conf. Sci. Inf. Technol., pp. 144–149, 2017.
- [13] S. M. H. Dadgar, M. S. Araghi, dan M. M. Farahani, "A novel text mining approach based on TF-IDF and support vector machine for news classification," Proc. 2nd IEEE Int. Conf. Eng. Technol. ICETECH 2016, no. March, pp. 112–116, 2016.
- [14] A. Chachra, P. Mehndiratta, dan M. Gupta, "Sentiment analysis of text using deep convolution neural networks," 2017 Tenth Int.

- Conf. Contemp. Comput., no. August, pp. 1–6, 2017.
- [15] P. Prameswari, “Mining Online Reviews in Indonesia ’ s Priority Tourist Destinations Using Sentiment Analysis and Text Summarization Approach,” no. iCAST, pp. 121–126, 2017.
- [16] H. Alam, “Sentiment Analysis for Bangla Sentences using Convolutional Neural Network,” pp. 22–24, 2017.
- [17] L.-S. Chen, C.-H. Liu, dan H.-J. Chiu, “A neural network based approach for sentiment classification in the blogosphere,” *J. Informetr.*, vol. 5, no. 2, pp. 313–322, 2011.
- [13] D. J. Suri dan K. K. Purnamasari, “Perbandingan Seleksi Fitur Untuk Klasifikasi Sentimen SVM Pada Twitter.”
- [14] J.A.Gliem dan R.R. Gliem, “Calculating, Interpreting, and Reporting Croanbach’s Alpha Reliability Coefficient for Likert-Type Scales,” p.87, 2003.