

SENTIMENT ANALYSIS TWITTER BAHASA INDONESIA BERBASIS WORD2VEC MENGGUNAKAN DEEP CONVOLUTIONAL NEURAL NETWORK

Hans Juwiantho^{*1}, Esther Irawati Setiawan¹², Joan Santoso¹², Mauridhi Hery Purnomo²³

¹Departemen Teknologi Informasi, Sekolah Tinggi Teknik Surabaya

²Departemen Teknik Elektro, Fakultas Teknologi Elektro, Institut Teknologi Sepuluh Nopember

³Departemen Teknik Komputer, Fakultas Teknologi Elektro, Institut Teknologi Sepuluh Nopember

Email: ¹hansjuwianthoo@gmail.com, ²esther@stts.edu, ³joan@stts.edu, ⁴hery@ee.its.ac.id

*Penulis Korespondensi

(Naskah masuk: 06 Februari 2019, diterima untuk diterbitkan: 14 Januari 2020)

Abstrak

Media sosial sebagai media informasi dan komunikasi mulai berkembang pesat sejak internet mudah diakses. Orang dengan mudah menyatakan pendapat, ekspresi, opini, dan informasi melalui tulisan pada media sosial. Opini atau informasi pada media sosial dapat digunakan untuk menilai baik atau buruk suatu *brand* perusahaan. Orang cenderung jujur dalam mengungkapkan perasaan terhadap sesuatu pada media sosial. Dengan menggunakan *sentiment analysis* terhadap opini dari pelanggan, analisis opini dapat dilakukan secara otomatis. Perusahaan dapat secara langsung mengetahui tingkat kepuasan pelanggan dan digunakan untuk meningkatkan kualitas pelayanan hingga menaikkan *brand* perusahaan. Penggunaan metode *classical machine learning* yang sudah banyak diterapkan pada *sentiment analysis*, tetapi metode tersebut tidak memperhatikan pentingnya urutan kata pada suatu kalimat. Metode *deep learning* dengan algoritme *Deep Convolutional Neural Network* ditawarkan untuk menjawab permasalahan tersebut dengan melakukan operasi *convolution* menggunakan *filter* sebesar ukuran *window* untuk mendapatkan fitur berdasarkan urutan kata. Model *Word2Vec* untuk Bahasa Indonesia digunakan sebagai representasi kata dalam bentuk vektor. Penggunaan *Word2Vec* juga mempercepat proses pelatihan dan meningkatkan akurasi algoritme *Deep Convolutional Neural Network*. Data yang digunakan dalam makalah ini adalah data *Twitter* Bahasa Indonesia dengan jumlah 999 *tweet*. Hasil percobaan yang telah dilakukan dengan algoritme *Deep Convolutional Neural Network* memiliki nilai akurasi terbaik sebesar 76,40%.

Kata kunci: *Sentiment Analysis, Word2Vec, Deep Convolutional Neural Network, Twitter* Bahasa Indonesia.

SENTIMENT ANALYSIS TWITTER INDONESIAN LANGUAGE WORD2VEC-BASED USING DEEP CONVOLUTIONAL NEURAL NETWORK

Abstract

Social media as information media and communication is growing rapidly since the internet is easily accessible. People easily express opinions, expressions, and information by writing on social media. Opinion or information on social media can be used to assess how good or bad a companies is. People tend to be honest in expressing feelings towards something on social media. With sentiment analysis, analysis of the opinions of customers can be done automatically. The company will know the level of customer satisfaction and can be used to improve the quality of service to raise the company's brand. The use of classical machine learning methods that have been widely applied to sentiment analysis ignoring the importance of the word order in a sentence. Deep Convolutional Neural Network algorithm is offered to answer these problems by carrying out convolution operations using filters as large as window size to get features based on word order. Word2Vec model for Indonesian is used as a word vector representation. The use of Word2Vec also reduce the training time and improve the accuracy of the Deep Convolutional Neural Network algorithm. The data used in this paper is Indonesian Twitter data with 999 tweets. The results of experiments that have been carried out with the Deep Convolutional Neural Network algorithm have the best accuracy value of 76.40%.

Keywords: *Sentiment Analysis, Word2Vec, Deep Convolutional Neural Network, Indonesian Twitter.*

1. PENDAHULUAN

Perkembangan teknologi digital terhadap media informasi dan komunikasi membuat masyarakat lebih

mudah mengakses dan membagikan informasi kepada orang lain tanpa dibatasi jarak. Salah satu alat komunikasi digital yang sering digunakan adalah

media sosial (Setiawan, 2017). Pada media sosial, orang dengan mudah dapat menyatakan pendapat, ekspresi, opini, dan informasi melalui tulisan dan gambar.

Media sosial berkembang dengan pesat semenjak internet mudah diakses. Salah satu media sosial populer di Indonesia adalah *Twitter* yang telah berdiri sejak 2006 dan semakin berkembang sampai sekarang. *Twitter* menjadi media sosial yang populer digunakan sebagai sumber data pada penelitian *sentiment analysis*. Di Indonesia, terdapat 29,4 juta pengguna *Twitter* dari total 328 juta pengguna di seluruh dunia (Abidin, Hasanuddin, & Mutiawani, 2017). Selain *Twitter*, pengguna media sosial *Facebook* di Indonesia mencapai urutan ketiga terbanyak pengguna *Facebook* di seluruh dunia (Statista.com, 2018). Dengan banyaknya pengguna aktif yang menggunakan media sosial terdapat juga banyak *review*, ulasan, dan komentar pada media sosial.

Perusahaan *startup* banyak yang melakukan ekstraksi *sentiment* dari data teks untuk membantu perkembangan perusahaan (Feldman, 2013). Dengan melakukan analisis terhadap opini dari pelanggan, perusahaan dapat mengukur kepuasan pelanggan dan meningkatkan kualitas pelayanan berdasarkan kritik dan masukan sehingga dapat menaikkan *brand* perusahaan. Penting bagi perusahaan untuk melakukan identifikasi *sentiment* secara otomatis terhadap setiap komentar mengenai perusahaan tersebut. Komentar dapat berupa komentar positif maupun komentar negatif. Identifikasi *sentiment* secara otomatis dapat dicapai dengan melakukan *sentiment analysis*.

Anastasia (2016) mengukur kepuasan pelanggan dari perusahaan “GOJEK” dan “GRAB” dengan membandingkan metode klasifikasi *Support Vector Machine*, *Naïve Bayes*, dan *Decision Tree*. Proses *pre-processing* yang dilakukan yaitu *retweet removal*, *case folding*, *stopwords filtering*, *tokenization*, dan *manual labeling*. Dengan total 9.191 *tweet* data dalam proses klasifikasi, diperoleh hasil bahwa “GRAB” memiliki nilai lebih tinggi daripada “GOJEK”. *Support Vector Machine* dan *Decision Tree* memberikan hasil akurasi terbaik sebesar 72,97%.

Penelitian lain mengenai *sentiment analysis* berfokus pada *brand* “GOJEK”. Penelitian tersebut menganalisis data *Twitter* yang berhubungan dengan “GOJEK” dan mengelompokkan data menjadi positif atau negatif dengan metode klasifikasi *Support Vector Machine* (SVM) dan TF-IDF sebagai metode ekstraksi fitur (Windasari, Uzzi, & Satoto, 2017). Penelitian dilakukan dengan melakukan teks *pre-processing* yaitu proses *convert emoticon*, *cleansing*, *stemming*, dan *convert negation*. Digunakan 1000 positif *tweet* dan 1.000 negatif *tweet* untuk proses *training* dengan 100 data untuk proses *testing* pada metode *Support Vector Machine* dan dihasilkan

akurasi dengan menggunakan *confusion matrix* sebesar 86%.

Somantri (2018) melakukan *sentiment analysis* untuk mengukur tingkat kepuasan pelanggan terhadap warung dan restoran kuliner di Kota Tegal dengan mengelompokkan data menjadi positif dan negatif yang selanjutnya diganti menjadi “Bagus” dan “Rata-rata”. Metode klasifikasi yang digunakan adalah *Support Vector Machine*. Pemilihan fitur menggunakan metode *Information Gain* dan *chi-square*. Penelitian dilakukan dengan mengolah teks menjadi *input* kedalam model yang digunakan. Proses pengolahan teks yang digunakan yaitu proses *transform cases*, *tokenize*, *filter tokens*, *stopword removal*, dan pembobotan dengan TF-IDF. Data yang digunakan adalah 80 dokumen teks, terbagi dalam 40 kategori “Rata-rata” dan 39 kategori “Bagus”. Hasil terbaik diperoleh pada pemilihan fitur *Information Gain* untuk model SVM dengan nilai akurasi sebesar 72,45%.

Naradhipa dan Purwarianti (2012) melakukan penelitian mengenai tantangan dalam melakukan *sentiment analysis* dalam Bahasa Indonesia di media sosial. Permasalahan dibagi menjadi dua, yaitu kata-kata dan kalimat. Kata-kata pada media sosial sering tidak memiliki struktur yang formal. Permasalahan pada kata-kata, seperti penyalahgunaan tanda baca, menjadikan angka sebagai huruf, penggunaan huruf yang diulang pada suatu kata, dan singkatan yang tidak sesuai standar. Permasalahan pada susunan kalimat dapat menyebabkan *sentiment* kalimat berubah apabila susunan kata tidak diperhatikan dan *tweet* dengan lebih dari satu kalimat dengan *sentiment* berbeda.

Kontribusi yang diberikan pada penelitian ini adalah diusulkannya algoritme *Deep Convolutional Neural Network* berbasis *Word2Vec* Bahasa Indonesia untuk *sentiment analysis* dalam Bahasa Indonesia. Metode yang ditawarkan memiliki kelebihan, yaitu pengenalan *sentiment* pada teks dengan memperhitungkan *sequence* kata pada kalimat, tidak seperti metode *classical* yang tidak memperdulikan urutan kata. *Word2Vec* Bahasa Indonesia digunakan sebagai inisialisasi kata menjadi vektor sehingga tidak dibutuhkan pelatihan *word* vektor dan pencarian fitur secara manual. *Word2Vec* akan menghasilkan vektor yang mirip pada kata yang memiliki makna yang mirip, hal ini akan mempermudah dalam pengenalan *sentiment analysis*. Penelitian ini menggunakan metode *deep learning* pada data berupa teks. Selanjutnya penelitian ini akan digabungkan dengan data gambar berdasarkan *tweet* yang memiliki teks dan gambar. Hal ini bertujuan untuk memperoleh *sentiment analysis* dengan tingkat akurasi yang lebih tinggi berdasarkan teks dan gambar.

2. SENTIMENT ANALYSIS

Sentiment analysis pada *Twitter* mulai dikenal pada tahun 2009 (Go, Bhayani, & Huang, 2009).

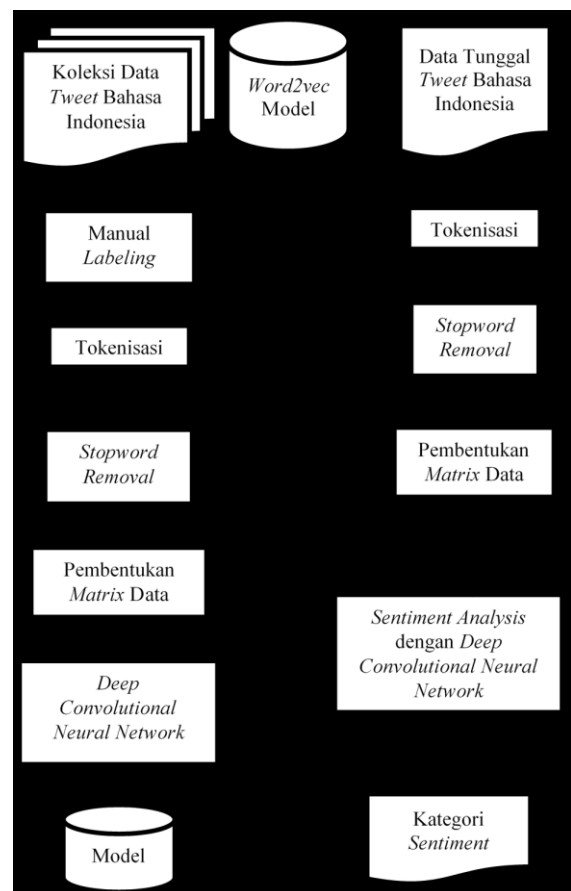
Twitter memberikan limit kepada setiap pengguna dengan maksimal 140 huruf untuk setiap *tweet*. Hal ini membuat proses klasifikasi *sentiment* terlihat mudah (Birmingham & Smeaton, 2010), tetapi sebenarnya menimbulkan permasalahan lain seperti singkatan yang tidak semestinya atau kata yang diganti dengan angka, seperti “57” yang seharusnya “setuju” (Finin et al., 2010). Media sosial tidak membatasi pengguna dalam penulisan opini, melainkan memberi kebebasan berekspresi. Hal ini membuat konten yang ditulis tidak memiliki struktur yang formal sehingga membuat proses *sentiment analysis* pada media sosial lebih susah. Dibutuhkan proses *pre-processing* untuk membersihkan data agar dapat lebih mudah dimengerti oleh algoritme *machine learning* (Abidin, Hasanuddin, & Mutiawani, 2017).

Penelitian yang sudah banyak dilakukan pada *sentiment analysis* dalam Bahasa Indonesia, menggunakan metode *classical machine learning*, seperti *Naïve Bayes* (NB), *Support Vector Machine* (SVM), *Decision Tree*, dan *Maximum Entropy* (ME) (Anastasia & Budi, 2016; Windasari, Uzzi & Satoto, 2017; Somantri et al., 2018), dengan melakukan pengenalan berdasarkan model *bag-of-words* yang tidak memperdulikan urutan dari suatu kata pada kalimat, lain halnya pada metode *deep learning* yang memperhitungkan urutan kata dalam kalimat (Vateekul & Koomsubha, 2016). Untuk Bahasa Inggris, sudah banyak penelitian yang menggunakan *deep learning* untuk *sentiment analysis* dan mendapatkan hasil yang jauh lebih baik dibandingkan metode *classical* (Vateekul & Koomsubha, 2016; Kim, 2014). Tetapi, belum ada penelitian yang mencoba melakukan *sentiment analysis* dengan metode *deep learning* pada Bahasa Indonesia. Dengan menggunakan metode *deep learning*, diharapkan pengenalan *sentiment* suatu kalimat lebih baik dibandingkan metode *classical*.

Penelitian yang sudah pernah dilakukan pada Bahasa Indonesia yang menggunakan *deep learning* untuk *sentiment analysis* menggunakan metode *Deep Neural Network* (Ilmania & Cahyawijaya, 2018). Berdasarkan percobaan yang dilakukan, ditemukan bahwa dengan menggunakan *word embedding* untuk membuat *similar matrix* dapat meningkatkan nilai akurasi. Percobaan dilakukan dengan menggunakan *word embedding* (*Word2Vec*) dengan ukuran 500 dimensi kata yang telah dilatih pada 27 juta kalimat dari data berita Bahasa Indonesia. Metode *Deep Neural Network* dibandingkan dengan metode yang sudah dilakukan pada penelitian sebelumnya yaitu *SVM* dan *Rule Based method* pada data yang sama dan aspek yang sudah ditentukan. Diperoleh hasil dari *Deep Neural Network* lebih baik 4 aspek dari total 7 aspek yang sudah dilakukan sebelumnya dengan metode *classical machine learning*.

Penelitian lain mengenai metode *deep learning* untuk *sentiment analysis* dilakukan dengan membandingkan algoritme *Long Short Term Memory* (LSTM) dan *Dynamic Convolutional Neural Network* (DCNN) dengan *classical machine learning* yaitu *Naïve Bayes*, *Support Vector Machine*, dan *Maximum Entropy* pada Bahasa Thailand dengan data yang didapatkan dari *Twitter* (Vateekul & Koomsubha, 2016). Hasil dari percobaan yang dilakukan pada Thailand *Twitter* data menunjukkan 2 algoritme *deep learning* lebih unggul dari metode *classical machine learning* dengan tingkat akurasi 75.35% pada metode DCNN.

Pada tahun 2014, penelitian mengenai *sentiment analysis* untuk Bahasa Inggris dengan menggunakan metode *deep learning* yaitu *Convolutional Neural Networks* (CNN) (Kim, 2014). Percobaan dilakukan pada beberapa *dataset*. Pada *dataset movie review* yang memiliki satu kalimat pada setiap *review* dengan *class sentiment* positif dan negative. Dilakukan percobaan pada 4 jenis CNN yaitu *CNN-rand*, *CNN-static*, *CNN-non-static*, dan *CNN-multichannel*. Diperoleh akurasi tertinggi pada model *CNN-non-static* dengan nilai 81,5%. *CNN-non-static* adalah CNN yang menggunakan *Word2Vec* model sebagai inialisasi kata dalam bentuk vektor sebelum digunakan dalam proses CNN. Metode CNN dengan



Gambar 1. Alur proses *sentiment analysis* dengan *Deep Convolutional Neural Network*

menggunakan CNN-*non-static* menjadi acuan dari percobaan yang akan dilakukan pada Bahasa Indonesia.

3. METODE PENELITIAN

Proses yang dilakukan pada penelitian ini dibagi menjadi tiga subproses yaitu *preprocessing* untuk mempersiapkan data, pembentukan matriks *data* sebagai input, dan pembentukan model *Deep Convolutional Neural Network* sebagai fitur klasifikasi. Urutan langkah setiap proses secara runtut diperlihatkan pada Gambar 1.

3.1. PREPROCESSING

Data *tweet* yang telah dikumpulkan dari *twitter*, banyak memiliki *noise* sehingga perlu adanya tahapan proses menghilangkan *noise* agar proses pengenalan *sentiment* menjadi lebih akurat dan dapat digunakan secara general. Dari total 3.219 *tweet* yang telah didapatkan, hanya 999 *tweet* yang digunakan pada proses berdasarkan Bahasa Indonesia sebagai acuan. Banyak *tweet* yang tidak dipakai karena bahasa yang digunakan bukan Bahasa Indonesia, tetapi cenderung Bahasa Inggris dan Malaysia. Sebagian besar data yang dikumpulkan mengandung kata “Asian games” karena saat pengumpulan data bertepatan dengan acara tersebut.

Proses akuisisi data *tweet* diperoleh dengan melakukan filter bahasa yaitu “Bahasa Indonesia” pada *website twitter* dan dilanjutkan dengan menelusuri halaman *website*. Kalimat dihasilkan dari setiap data *tweet* dibantu sebuah *crawler* sederhana berdasarkan *rule*. Untuk mendapatkan *data* yang bersih dan siap digunakan, dilakukan beberapa proses sebagai berikut:

- *Manual Labeling*: Proses *manual labelling* dilakukan dengan memberikan *sentiment* secara manual untuk setiap *tweet*. *Sentiment* dibagi menjadi 3 pilihan, yaitu positif, negatif, dan netral. Data *tweet* yang sudah diberi label digunakan untuk melatih *classifier*.
- *Tokenisasi*: Proses tokenisasi dilakukan dengan memisahkan kumpulan kata pada *tweet* menjadi kata yang berdiri sendiri dan diberi *token* untuk setiap kata.
- *Stopword Removal*: Proses ini dilakukan untuk menghilangkan kata-kata yang tidak penting karena seringnya muncul disetiap *tweet* sehingga dapat mengganggu dalam proses pengenalan *sentiment*. Daftar *stopword* yang digunakan diperoleh dari Librian (2017). Proses ini dilengkapi dengan menghilangkan *hyperlink*, *case folding* untuk mengubah seluruh kata menjadi seragam dalam huruf kecil, dan menghilangkan tanda baca dengan *regular expression*.

3.2. PEMBENTUKAN MATRIKS

Algoritme *Deep Convolutional Neural Network* merupakan sebuah model jaringan yang menerima

masukan berbentuk matriks. Penerapan pada data yang berupa gambar dilakukan dengan memanfaatkan angka *pixel* pada gambar sebagai isi dari matriks. Hal ini berbeda dengan data teks yang tidak memiliki angka sehingga perlu dilakukan *convert* menjadi angka.

Kim (2014) melakukan penelitian dengan membandingkan inisialisasi angka yang digunakan untuk data teks pada metode *Convolutional Neural Network*. Akurasi dengan menggunakan *pre-trained word vectors* jauh lebih tinggi dibandingkan dengan inisialisasi dengan angka secara acak. *Word vector* digunakan untuk merepresentasikan teks menjadi vektor dari hasil pembelajaran dengan algoritme *Neural Network*. Setiap kata akan memiliki vektor yang mewakili makna dari kata tersebut. Kata yang memiliki kemiripan makna akan menghasilkan vektor dengan angka yang hampir sama. Proses merubah teks menjadi angka dilakukan dengan menggunakan *pre-trained word vector* atau *Word2Vec* Bahasa Indonesia yang diperoleh dari Santoso (2018).

Dimensi vektor yang digunakan berdasarkan *Word2Vec* adalah 50, 100, 150, 200, 250, dan 300. Penggunaan variasi dimensi bertujuan untuk mendapatkan dimensi dengan ukuran yang dapat menghasilkan akurasi terbaik. Semakin kecil dimensi yang digunakan maka semakin banyak informasi yang dibuang. Hal ini dapat membantu dalam proses pengenalan *sentiment*.

Setiap kata pada kalimat dirubah menjadi vektor sesuai *Word2Vec* dan setiap kalimat ditambahkan *padding* untuk menyesuaikan kalimat yang panjang dan pendek menjadi ukuran yang seragam sehingga matriks input menjadi sama untuk setiap kalimat. *Padding* yang digunakan adalah *post padding*, yaitu menambahkan vektor kosong pada akhir kata hingga ukuran kalimat menjadi sama dengan kalimat terpanjang pada keseluruhan data. Setelah proses *convert* kata dan *padding* dilakukan, dihasilkan matriks data yang dapat digunakan sebagai input pada algoritme *Deep Convolutional Neural Network*.

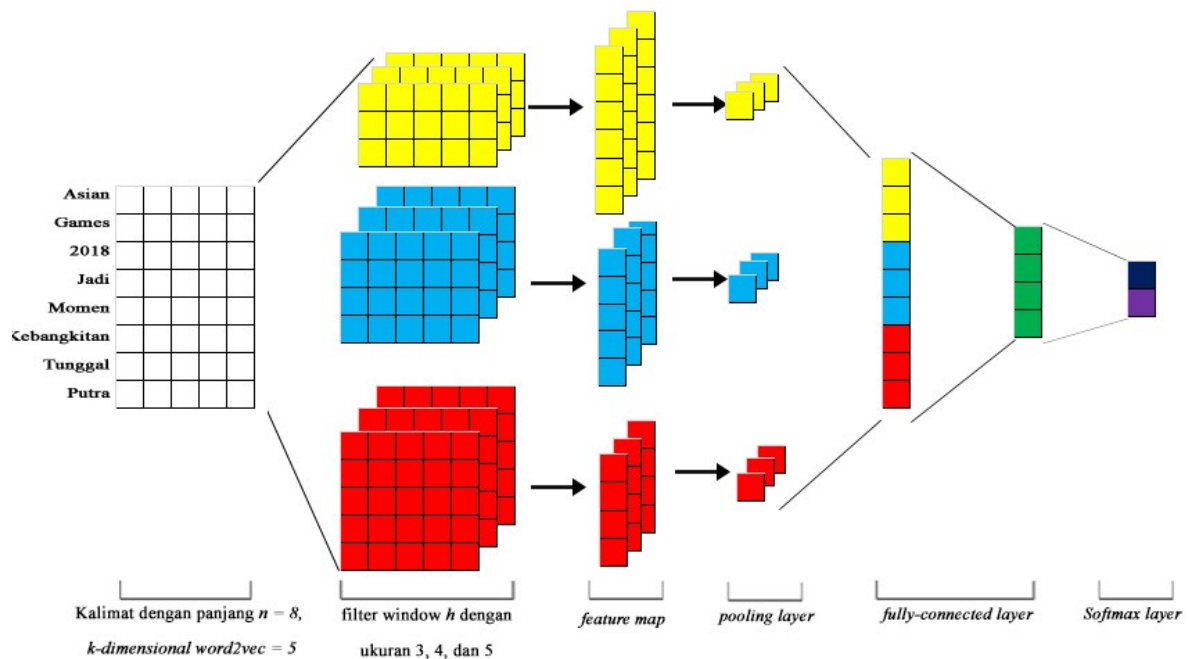
3.3. PEMBENTUKAN MODEL

Arsitektur *Deep Convolutional Neural Network* yang digunakan pada makalah ini mengacu pada arsitektur yang diperoleh dari Kim (2014) dengan tambahan modifikasi. Model *Deep Convolutional Neural Network* yang digunakan pada makalah ini dapat dilihat pada Gambar 2.

Jika $x_i \in \mathbb{R}^k$ merepresentasikan *k-dimensional Word2Vec* berdasarkan kata ke *i* pada kalimat, kalimat dengan panjang *n* (ditambahkan *post padding* jika diperlukan) dituliskan menjadi fungsi pada persamaan (1).

$$x_{1:n} = x_1 \oplus x_2 \oplus \dots \oplus x_n \quad (1)$$

Operator \oplus merupakan operator penggabungan dan digunakan untuk menggabungkan kata yang telah



Gambar 2. Model Deep Convolutional Neural Network

dirubah menjadi vektor menjadi bentuk matriks. Operasi *Convolution* menggunakan filter $w \in \mathbb{R}^{hk}$, dengan ukuran *window* h sebagai banyaknya jumlah kata urut yang diproses. Ukuran *window* yang digunakan pada makalah ini adalah 3, 4, dan 5. Fitur c_i dihasilkan dari *window* berukuran h dilakukan *dot product* dengan matriks input $x_{i:i+h-1}$ dan dilanjutkan dengan fungsi aktivasi *rectifier linear unit*. Fungsi fitur dituliskan pada persamaan (2).

$$c_i = \max(0, w \cdot x_{i:i+h-1}) \quad (2)$$

Filter diterapkan kepada setiap *window* kata dalam suatu kalimat $\{x_{1:h}, x_{2:h+1}, \dots, x_{n-h+1:n}\}$ untuk menghasilkan *feature map* dengan $c \in \mathbb{R}^{n-h+1}$. Fungsi *feature map* dituliskan pada persamaan (3).

$$c = [c_1, c_2, \dots, c_{n-h+1}] \quad (3)$$

Untuk menghindari model yang terlalu cocok pada data pelatihan atau *overfitting*, dibutuhkan *dropout layer* yang berfungsi memilih secara acak *neuron* untuk tidak dipakai selama proses pelatihan (Srivastava et al., 2014). Probabilitas *dropout layer* yang digunakan pada penelitian ini sebesar 0,5. Proses pelatihan dilanjutkan dengan melakukan *pooling* pada *feature map*. Metode *pooling* yang digunakan adalah *MaxPooling*, yaitu mengambil nilai maksimum $\hat{c} = \max\{c\}$ sebagai fitur berdasarkan suatu filter. Tujuan proses *pooling* adalah untuk mendapatkan fitur terpenting yang mewakili fitur lain untuk setiap *feature map*.

Fitur yang telah didapatkan dari hasil *pooling* digunakan untuk proses klasifikasi pada *fully-connected layer* dan dilatih dengan algoritme

Backpropagation (Hecht-Nielsen, 1988). Setiap fitur akan disambungkan dengan setiap *neuron* pada *fully-connected layer* dengan bobot independen dan jumlah *neuron* sebanyak 50. Setiap *neuron* dari *fully connected layer* menjadi input bagi *softmax layer*. Jumlah *neuron* pada *softmax layer* berdasarkan

```

Function AdamOptimizer ( $\alpha$ : Learning Rate,
 $\beta_1 = 0.9, \beta_2 = 0.999, \theta_0$ : Initial Parameter
Vector)
1.  $m_0 = 0$ 
2.  $v_0 = 0$ 
3.  $t = 0$ 
4. WHILE  $\theta_t$  not converged DO
5.    $t = t + 1$ 
6.    $g_t = \nabla_{\theta} f_t(\theta_{t-1})$ 
7.    $m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$ 
8.    $v_t = \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$ 
9.    $\hat{m}_t = m_t / (1 - \beta_1^t)$ 
10.   $\hat{v}_t = v_t / (1 - \beta_2^t)$ 
11.   $\theta_t = \theta_{t-1} - \alpha \cdot \hat{m}_t / (\sqrt{\hat{v}_t} + \epsilon)$ 
12. END WHILE
13. RETURN  $\theta$ ,
    
```

banyaknya sentiment yang digunakan, 2 *neuron* untuk 2 *sentiment* dan 3 *neuron* untuk 3 *sentiment*. Fungsi *softmax* digunakan untuk mendapatkan nilai

Gambar 3. Kerangka Algoritme Adam Optimizer

probabilitas yang telah dilakukan normalisasi pada semua *class*.

Dalam melakukan training, dibutuhkan *optimizer* untuk mengurangi *loss function* atau kesalahan yang terdapat pada *neuron* dan filter. *Optimizer* dilakukan dengan *update* terhadap *neuron*

dan filter pada setiap iterasi. Pada penelitian ini, *Adam optimizer* digunakan dalam melakukan training *Deep Convolutional Neural Network* (Kingma & Ba, 2014). Kerangka algoritme *Adam optimizer* diperlihatkan pada Gambar 3. Algoritme *Adam optimizer* dilakukan dengan menghitung rata-rata secara *exponential* dari *gradient* dan pangkat *gradient*. Parameter Algoritme *Adam Optimizer* yang digunakan pada penelitian ini adalah *learning rate* = 0,001, $\beta_1 = 0,9$, dan $\beta_2 = 0,999$.

4. HASIL DAN PEMBAHASAN

Dalam bagian ini, dijelaskan tentang uji coba yang dilakukan untuk mendapatkan model yang memiliki akurasi terbaik dengan mengganti parameter yang digunakan. Parameter yang diteliti yaitu ukuran dimensi *Word2Vec*, ukuran filter, dan jumlah label *sentiment*. Alat ukur yang digunakan adalah akurasi yang dituliskan pada persamaan (4). *True Positive* merupakan jumlah tebakan yang benar dan Total Data adalah jumlah data yang diuji.

$$\text{Akurasi} = \frac{\text{True Positive}}{\text{Total Data}} \quad (4)$$

Data yang digunakan dibagi menjadi *training data* dan *testing data*, dengan pembagian 75% untuk *training data* dan 25% untuk *testing data*. Dimensi *Word2Vec* dan ukuran filter terbaik dilihat dengan menggunakan algoritme *Deep Convolutional Neural Network* pada pengujian terhadap *testing data*. Ukuran untuk dimensi *Word2Vec* yang digunakan adalah 50, 100, 150, 200, 250, dan 300. Jumlah filter yang dicoba adalah 50 dan 100. Jumlah label *sentiment* yang ditentukan adalah 2 dan 3. Pada 2 *sentiment*, data dibagi menjadi data positif atau tidak positif dan untuk 3 *sentiment* terdapat positif, netral, dan negatif. Percobaan dengan 2 *sentiment* dilakukan untuk mengetahui pengaruh keseimbangan data terhadap akurasi yang dihasilkan. Percobaan dengan metode *classical* dilakukan dengan membandingkan metode *Deep Convolutional Neural Network*, *Naïve Bayes*, dan *Support Vector Machine*. Statistik perbandingan data terhadap label *sentiment* terdapat pada Tabel 1. Hasil percobaan dengan 3 *sentiment* dapat dilihat pada Tabel 2 dan Tabel 3, sedangkan hasil percobaan dengan 2 *sentiment* dapat dilihat pada Tabel 4. Hasil perbandingan metode *classical machine learning* dan metode *Deep Convolutional Neural Network* pada 2 *sentiment* dataset dan 3 *sentiment* dataset dapat dilihat pada Tabel 5.

Tabel 1. Perbandingan Data Terhadap label Sentiment

| Label | Jumlah Data |
|---------|-------------|
| Positif | 419 |
| Negatif | 453 |
| Neutral | 127 |

Hasil terbaik percobaan pada filter ukuran 50 di Table 2 diperoleh pada *Word2Vec* dimensi 100 dengan nilai akurasi 69,20% diikuti dimensi 200

dengan nilai akurasi 66%. Rata-rata nilai akurasi dari hasil percobaan yang dilakukan dengan filter ukuran 50 adalah 65,43%.

Tabel 2. Hasil Percobaan Word2Vec untuk ukuran Filter 50 dengan 3 Label Sentiment

| No. | Filter | Dimensi | Akurasi |
|-----------|--------|---------|---------|
| 1 | 50 | 50 | 63,10% |
| 2 | 50 | 100 | 69,20% |
| 3 | 50 | 150 | 63,90% |
| 4 | 50 | 200 | 66% |
| 5 | 50 | 250 | 65,60% |
| 6 | 50 | 300 | 64,80% |
| Rata-rata | | | 65,43% |

Percobaan dengan filter ukuran 100 di Tabel 3 mendapatkan akurasi terbaik pada *Word2Vec* dimensi 100 dengan akurasi 65,60%. Nilai akurasi rata-rata dari seluruh percobaan dengan filter ukuran 100 berada pada angka 64,32%.

Tabel 3. Hasil Percobaan Word2Vec untuk ukuran Filter 100 dengan 3 Label Sentiment

| No. | Filter | Dimensi | Akurasi |
|-----------|--------|---------|---------|
| 1 | 100 | 100 | 65,60% |
| 2 | 100 | 150 | 64,30% |
| 3 | 100 | 200 | 63,10% |
| 4 | 100 | 250 | 63,90% |
| 5 | 100 | 300 | 64,70% |
| Rata-rata | | | 64,32% |

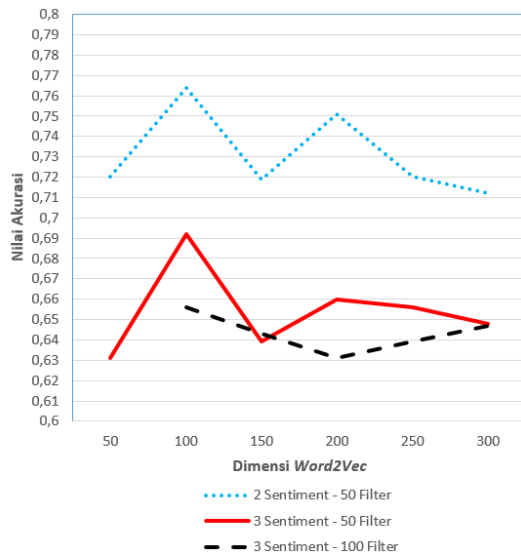
Percobaan terakhir dilakukan pada 2 *sentiment label* dan filter ukuran 50 di Tabel 4 menghasilkan kinerja terbaik pada *Word2Vec* dimensi 100 dengan nilai akurasi 76,40% dan diikuti dimensi 200 dengan nilai akurasi 75,10%. Rata-rata nilai akurasi dari hasil percobaan yang dilakukan pada 2 *sentiment* label dan filter ukuran 50 adalah 73,10%.

Tabel 4. Hasil Percobaan Word2Vec untuk ukuran Filter 50 dengan 2 Label Sentiment

| No. | Filter | Dimensi | Akurasi |
|-----------|--------|---------|---------|
| 1 | 50 | 50 | 72% |
| 2 | 50 | 100 | 76,40% |
| 3 | 50 | 150 | 71,90% |
| 4 | 50 | 200 | 75,10% |
| 5 | 50 | 250 | 72% |
| 6 | 50 | 300 | 71,20% |
| Rata-rata | | | 73,10% |

Nilai rata-rata dari percobaan dimensi *Word2Vec* terbaik diperoleh pada filter ukuran 50. Nilai rata-rata akurasi pada 3 label *sentiment* sebesar 65,43% dan pada 2 label *sentiment* sebesar 73,10%. Dapat disimpulkan bahwa penggunaan data yang *balance* pada 2 label *sentiment* menghasilkan nilai akurasi lebih baik dibandingkan dengan data yang tidak *balance*. Ukuran dimensi *Word2Vec* terbaik dari hasil setiap percobaan adalah sebesar 100. Hasil

percobaan ditampilkan menggunakan grafik pada Gambar 4.



Gambar 4. Grafik nilai akurasi untuk algoritme *Deep Convolutional Neural Network*

Percobaan metode terhadap dataset dilakukan untuk melihat nilai akurasi dari metode *classical machine learning* dengan metode *deep learning* yang diusulkan. Metode yang diuji yaitu *Naïve bayes*, SVM dengan Bag-of-words, SVM dengan TFIDF, dan DCNN. Hasil akurasi perbandingan metode dapat dilihat pada Tabel 5. Metode *Naïve bayes* dan SVM mendapatkan nilai akurasi lebih kecil dibandingkan metode *deep learning* terhadap dataset yang sama. Tingkat akurasi dari metode NB dan SVM sebesar 70,90% dan 70,49% untuk dataset 2 sentiment, sedangkan pada dataset 3 sentiment dihasilkan akurasi 59,83% pada kedua metode. Nilai akurasi tertinggi diperoleh pada metode *Deep Convolutional Neural Network* dengan nilai akurasi yang jauh lebih tinggi dibandingkan dengan metode *classical machine learning* yaitu 76,40% pada 2 sentiment dan 69,20% pada 3 sentiment.

| Metode | 2 Sentiment | 3 Sentiment |
|-------------|-------------|-------------|
| NB | 70,49% | 59,83% |
| NB + TFIDF | 70,90% | 58,19% |
| SVM + BOW | 69,67% | 59,02% |
| SVM + TFIDF | 70,49% | 59,83% |
| DCNN | 76,40% | 69,20% |

Perbandingan nilai akurasi terhadap penelitian sebelumnya dengan menggunakan metode *classical deep learning*, seperti metode SVM, NB, *Decision tree*, dan *rule-based* (Naradhupa & Purwarianti, 2012; Anastasia & Budi, 2016; Windasari, Uzzi, & Satoto, 2017; Somantri et al., 2018). Pada penelitian yang dilakukan oleh Anastasia (2016) dengan jumlah data 9,191 dengan 2 *sentiment* label, diperoleh nilai akurasi tertinggi 72,97% dengan menggunakan SVM dan *Decision Tree*. Hasil kinerja akurasi dengan

metode *classical* lebih kecil 3,43% dibandingkan dengan algoritme *deep learning* yang diusulkan pada makalah ini.

Penelitian terkait lainnya dengan metode *classical* dilakukan oleh Windasari (2017), Naradhupa (2012), dan Somantri (2018), mendapatkan hasil akurasi terbaik dengan metode SVM sebesar 86%, 83,5%, dan 72,45% dengan jumlah data yang digunakan sebanyak 100, 120, dan 80. Percobaan algoritme *Deep Convolutional Neural Network* yang diusulkan menggunakan jumlah data hampir sepuluh kali lebih besar jika dibandingkan jumlah data dari percobaan sebelumnya pada metode *classical*. Proses pengenalan fitur di penelitian ini dilakukan secara otomatis menggunakan filter pada metode *Deep Convolutional Neural Network* tanpa harus dilakukan secara manual dan *Word2Vec* Bahasa Indonesia digunakan sebagai inialisasi kata dalam bentuk vektor yang sudah dilatih sebelumnya.

5. KESIMPULAN

Penggunaan model *Word2Vec* dengan dimensi 100 sebagai inialisasi memberikan nilai akurasi terbaik. Hal ini dapat disebabkan karena dimensi dengan ukuran 100 cocok untuk mewakili *sentiment* dari data yang digunakan. Ukuran dimensi yang terlalu besar menyebabkan munculnya vektor tambahan tetapi tidak mewakili *sentiment*. Sebaliknya pada dimensi kecil, makna dari suatu vektor yang digunakan kurang menunjukkan makna *sentiment*.

Pemilihan jumlah filter yang sesuai juga mempengaruhi dari algoritme *Deep Convolutional Neural Network*. Hasil terbaik dari algoritme *Deep Convolutional Neural Network* diperoleh pada jumlah filter ukuran 50 dengan rata-rata nilai akurasi 65,43% pada 3 *sentiment* label dan 73,10% pada 2 *sentiment* label. Terlalu banyak filter yang digunakan akan menghasilkan fitur-fitur kurang penting dan dapat menjadi *noise*.

Perbandingan metode *classical* dan *deep learning* diperoleh hasil akurasi pada metode *classical* sebesar 70,90% pada 2 *sentiment* dan 59,83% pada 3 *sentiment*. Akurasi terbaik pada metode *deep learning* sebesar 76,40% pada 2 *sentiment* dan 69,20% pada 3 *sentiment*. Hasil dari metode *deep learning* lebih baik 5,5% pada 2 *sentiment* dan 9,37% pada 3 *sentiment* dibandingkan dengan metode *classical machine learning*.

Pada penelitian mendatang, akan dicoba untuk menerapkan algoritme *Deep Convolutional Neural Network* pada data gambar beserta tulisan sebagai input untuk klasifikasi *sentiment*. Upaya lain yang dilakukan dalam penelitian mendatang ialah dengan mengumpulkan data yang lebih banyak.

6. DAFTAR PUSTAKA

ABIDIN, T.F., HASANUDDIN, M. dan MUTIAWANI, V., 2017. N-grams based

- features for Indonesian tweets classification problems. *Proceedings - 2017 International Conference on Electrical Engineering and Informatics: Advancing Knowledge, Research, and Technology for Humanity, ICELTICs 2017*, 2017–Octob(ICELTICs), hal.307–310.
- ANASTASIA, S. dan BUDI, I., 2016. Twitter sentiment analysis of online transportation service providers. *2016 International Conference on Advanced Computer Science and Information Systems (ICACIS)*, hal.359–365.
- BERMINGHAM, A. dan SMEATON, A.F., 2010. Classifying Sentiment in Microblogs: Is Brevity an Advantage? *Proceedings of the 19th ACM international conference on Information and knowledge management - CIKM '10*, hal.1833.
- FELDMAN, R., 2013. Techniques and Applications for Sentiment Analysis. *Commun. ACM*, [Online] 56(4), hal.82–89. Tersedia pada: <<http://doi.acm.org/10.1145/2436256.2436274>>.
- FININ, T., MURNANE, W., KARANDIKAR, A., KELLER, N., MARTINEAU, J. dan DREDZE, M., 2010. Annotating named entities in Twitter data with crowdsourcing. *Proceedings of NAACL-HLT*, [Online] 2010(January), hal.80–88. Tersedia pada: <<http://dl.acm.org/citation.cfm?id=1866696.1866709>>.
- GO, A., BHAYANI, R. dan HUANG, L., 2009. Twitter Sentiment Classification using Distant Supervision. *CS224N Project Report*, 1, hal.12.
- HECHT-NIELSEN, R., 1988. Theory of the backpropagation neural network. *International 1989 Joint Conference on Neural Networks*, hal.593–605 vol.1.
- ILMANIA, A. dan CAHYAWIJAYA, S., 2018. Aspect Detection and Sentiment Classification using Deep Neural Network for Indonesian Aspect-Based Sentiment Analysis. *2018 International Conference on Asian Language Processing (IALP)*, hal.62–67.
- KIM, Y., 2014. Convolutional Neural Networks for Sentence Classification. *CoRR*, [Online] abs/1408.5. Tersedia pada: <<http://arxiv.org/abs/1408.5882>>.
- KINGMA, D.P. dan BA, J., 2014. Adam: A Method for Stochastic Optimization. [Online] hal.1–15. Tersedia pada: <<http://arxiv.org/abs/1412.6980>>.
- LIBRIAN, A., 2017. *High quality stemmer library for Indonesian Language (bahasa)*, *GitHub*. [Online] Tersedia pada: <<https://github.com/sastrawi/sastrawi>> [Diakses 18 Apr 2018].
- NARADHIPA, A.R. dan PURWARIANTI, A., 2012. Sentiment Classification for Indonesian Messages in Social Media. *International Conference on Electrical Engineering and Informatics*, (July), hal.2–5.
- SANTOSO, J., SOETIONO, A.D.B., GUNAWAN, G., SETYATI, E., YUNIARNO, E.M., HARIADI, M. dan PURNOMO, M.H., 2018. Self-Training Naive Bayes Berbasis Word2Vec untuk Kategorisasi Berita Bahasa Indonesia. *JNTETI*, 7(2), hal.158–166.
- SETIAWAN, W., 2017. Era Digital dan Tantangannya. *Seminar Nasional Pendidikan 2017, Sukabumi*, [Online] hal.1–9. Tersedia pada: <<http://eprints.ummi.ac.id/151/>>
- SOMANTRI, O., APRILIANI, D., INFORMATIKA, J.T., HARAPAN, P. dan TEGAL, B., 2018. Support vector machine berbasis feature selection untuk Sentiment analysis kepuasan pelanggan terhadap pelayanan warung dan restoran kuliner kota tegal. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 5(5), hal.537–548.
- SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., SUTSKEVER, I. dan SALAKHUTDINOV, R., 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, [Online] 15, hal.1929–1958. Tersedia pada: <<http://jmlr.org/papers/v15/srivastava14a.html>>.
- STATISTA.COM, 2018. *Leading countries based on number of Facebook users as of October 2018 (in millions)*. [Online] Tersedia pada: <<https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/>> [Diakses 28 Nov 2018].
- VATEEKUL, P. dan KOOMSUBHA, T., 2016. A study of sentiment analysis using deep learning techniques on Thai Twitter data. *2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, [Online] hal.1–6. Tersedia pada: <<http://ieeexplore.ieee.org/document/7748849/>>.
- WINDASARI, I.P., UZZI, F.N. dan SATOTO, K.I., 2017. Sentiment analysis on Twitter posts: An analysis of positive or negative opinion on GoJek. *2017 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, hal.266–269.