

Deteksi Depresi Berdasarkan Data Twitter Menggunakan *Convolutional Neural Network*

Faris Salbari, Dini Adni Navastara, Nanik Suciati
Departemen Informatika, Fakultas Teknologi Informasi dan Komunikasi,

Institut Teknologi Sepuluh Nopember (ITS)
Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia
e-mail: dini_navastara@if.its.ac.id

Abstrak—Twitter adalah layanan jejaring sosial yang memiliki pengguna cukup banyak di Indonesia. Karena Twitter berbentuk *microblogging*, pengguna banyak memanfaatkannya untuk bercerita tentang masalah di kehidupan sehari-harinya. Dari tweet ini dapat dideteksi apakah pengguna ada indikasi depresi atau tidak.

Data yang digunakan pada penelitian ini diambil dengan metode *webcrawling* dan didapatkan 1000 *tweet* Twitter yang kemudian dilabeli secara manual menjadi 500 *tweet* depresi dan 500 *tweet* bukan depresi. Kemudian pada data tersebut dilakukan *text preprocessing*. Diterapkan juga *word embedding* dengan metode *Word2Vec Skip-gram* dan *word encoding* seperti *sequencing* dan *padding* agar data dapat diproses oleh algoritma CNN. Selanjutnya proses *training* data menggunakan model-model CNN yaitu model AlexNet, ZFNet, VGG16, VGG19, LeNet-5, serta model yang diusulkan dan menggunakan *Stratified 5-Fold cross validation* sebagai metode validasi.

Hasil uji coba menunjukkan bahwa nilai rata-rata akurasi tertinggi sebesar 75,33% berdasarkan arsitektur model yang diusulkan. Arsitektur model tersebut terdiri dari 3 *layer* dengan jumlah *filter* sebanyak 16 pada tiap *layer* dan menggunakan *optimizer* Nadam dengan *learning rate* 0,0001. Arsitektur model tersebut memiliki rata-rata akurasi yang lebih tinggi dari lima arsitektur model lainnya.

Kata Kunci— *convolutional neural network, deteksi depresi, text preprocessing, twitter, word embedding, word encoding*

I. PENDAHULUAN

Twitter adalah sebuah jejaring sosial yang hingga saat ini masih banyak digunakan oleh masyarakat Indonesia. Terbukti hingga Januari 2018, 27% dari pengguna jejaring sosial di Indonesia yang berjumlah 130 juta pengguna memiliki akun Twitter [1]. Twitter masih digunakan karena fiturnya yang memungkinkan pengguna untuk dapat bercerita dan berbagi pengalaman pada konten yang disebut “*tweet*” kepada pengguna lainnya.

Analisa isi dari *tweet* dewasa ini sudah menjadi cara yang cukup sering digunakan untuk memahami dan memprediksi perilaku sosial manusia. Karena frekuensi penggunaan yang cukup tinggi dan pengguna tersebar di masyarakat luas,

Twitter menjadi sumber yang cukup kaya untuk memberikan data yang dapat digunakan untuk menganalisis berbagai macam perilaku tersebut.

Dewasa ini banyak berita kematian dengan cara bunuh diri. Menurut data statistik WHO, pada tahun 2014 sebanyak 23% dari jumlah kematian di dunia diakibatkan oleh gangguan jiwa dan penyalahgunaan obat-obatan. Kemudian menurut data statistik yang dikeluarkan oleh situs *World Population Review*, tiga dari seratus ribu penduduk Indonesia melakukan bunuh diri pada tahun 2018 [2].

Tahun 1980, Yann LeCun menemukan sebuah metode untuk melakukan pengenalan obyek gambar, metode ini kemudian disebut *Convolutional Neural Network* atau CNN [3]. Walaupun metode ini lebih terkenal untuk melakukan klasifikasi obyek pada gambar, dewasa ini metode CNN sudah mulai digunakan untuk klasifikasi *Natural Language Processing*, seperti *Sentiment Analysis* dan *Spam Detection*.

Dalam makalah ini, dilakukan pengambilan data teks dari jejaring sosial Twitter menggunakan *crawler*, untuk kemudian diterapkan ke sebuah model CNN untuk mendeteksi depresi berdasarkan data Twitter. Diharapkan model yang dibangun dapat mendeteksi secara akurat.

II. ANALISIS DAN PERANCANGAN

A. Twitter

Twitter merupakan sebuah layanan jejaring sosial yang memungkinkan penggunanya untuk mengirim dan membaca pesan berbasis teks yang disebut “*tweet*”. *Tweet* pada awalnya memiliki batas sebanyak 140 karakter. Namun, pada 7 November 2017 batasan tersebut dinaikkan hingga 280 karakter untuk semua Bahasa kecuali Bahasa Jepang, Korea, dan Mandarin.

Twitter didirikan pada Maret 2006 oleh Jack Dorsey, Noah Glass, Biz Stone, serta Evan Williams dan situsnya diluncurkan pada Juli 2006. Pada tahun 2012, lebih dari 100 juta pengguna mengirim *tweet* sebanyak 340 juta tiap harinya dan Twitter menangani 1,6 miliar pencarian. Pada tahun 2013, Twitter menjadi salah satu dari sepuluh situs yang paling sering dikunjungi. Twitter disebut juga sebagai “*the SMS of*

Internet". Tercatat hingga awal tahun 2018, Twitter memiliki 336 juta pengguna aktif.

Pengguna yang sudah terdaftar dapat mengirim dan membaca *tweet*. Sedangkan pengguna yang tidak terdaftar hanya dapat membaca *tweet* yang dikirim pengguna lain. Pengguna dapat mengakses Twitter melalui situs Twitter, pesan singkat (SMS), atau melalui aplikasi pada perangkat seluler.

Twitter memiliki fitur untuk saling mengikuti (*follow*) antar penggunanya yang memungkinkan *tweet* pengguna yang diikuti ditampilkan oleh sistem pada lini masa para pengikut pengguna tersebut. Fitur *retweet*, *reply*, *like*, serta *direct message* juga tersedia sebagai sarana untuk berhubungan antar pengguna.

Beberapa fitur utama lainnya yang dimiliki Twitter adalah *mention*, *hashtag*, dan *trending topic*. *Mention* adalah fitur yang memungkinkan pengguna untuk membalas ataupun mengirim *tweet* pengguna lain dengan cara menggunakan tanda "@" yang diikuti nama pengguna. Fitur *hashtag* adalah fitur yang memudahkan pengguna untuk mengelompokkan dan mencari *tweet* berdasarkan topik bahasan dari *tweet* tersebut dengan cara menggunakan tanda "#" yang diikuti kata atau frasa topik yang diinginkan. Kemudian kata atau frasa topik yang sedang hangat dibicarakan akan ditampilkan pada fitur *trending topic*. Fitur ini membantu Twitter dan penggunanya untuk memahami apa yang sedang terjadi di dunia.

B. Depresi

Depresi adalah gangguan jiwa yang umum dan dapat memengaruhi diri secara negatif, baik dalam hal perasaan, pikiran, dan tindakan. Depresi dapat menyebabkan perasaan sedih dan/atau kehilangan minat dari suatu aktivitas yang sebelumnya disukai. Depresi juga dapat menyebabkan bermacam masalah emosi dan fisik serta dapat mengurangi kemampuan beraktivitas dalam pekerjaan ataupun di rumah.

Gejala depresi dapat bermacam-macam, contohnya

- Merasakan sedih terus menerus
- Kehilangan minat dan kesenangan saat melakukan hal yang sebelumnya dinikmati
- Perubahan nafsu makan yang menyebabkan turun atau naiknya berat badan tidak teratur yang bukan disebabkan oleh diet
- Kurang waktu tidur atau tidur terlalu lama yang menyebabkan waktu tidur tidak teratur
- Mudah kehilangan energi dan lelah
- Sering melakukan kegiatan yang tidak bertujuan
- Sering merasa salah dan tidak berguna
- Sulit untuk berpikir, berkonsentrasi, dan memilih keputusan
- Memiliki pikiran untuk mati atau bunuh diri

Depresi dapat menimpa siapapun, bahkan seseorang yang menurutnya sudah tinggal di lingkungan yang ideal. Ada beberapa faktor yang dapat memengaruhi depresi, di antaranya adalah faktor kepribadian dan lingkungan. Orang dengan tingkat kepercayaan diri yang rendah dan mudah mengalami stress serta orang yang bersifat pesimis adalah orang-orang

yang mudah mengalami depresi. Sering menerima tindak kekerasan, ditolak oleh lingkungan, dan mengalami kemiskinan juga dapat membuat orang menjadi rentan terhadap depresi [4].

C. Convolutional Neural Network

Convolutional Neural Network (CNN) adalah salah satu jenis *neural network* yang biasa digunakan pada data gambar, seperti untuk mendeteksi dan mengenali obyek pada sebuah gambar. Namun CNN juga dapat digunakan pada data teks, dan sudah teruji memiliki akurasi yang cukup baik. Secara garis besar, CNN tidak jauh berbeda dengan *neural network* biasanya. CNN terdiri dari *neuron* yang memiliki *weight*, *bias*, dan *activation function* yang bersifat *non-linear*.

Pada sebuah CNN umumnya terdapat *convolutional layer*, *pooling layer*, dan *fully connected layer* [5].

1. Convolution Layer

Lapisan ini adalah inti dari CNN yang melakukan sebagian besar dari komputasi. Parameter dari lapisan ini berisi *filter*. *Filter* tersebut akan dilakukan komputasi *dot product* dengan *input*. Kemudian komputasi tersebut akan menghasilkan fungsi aktivasi 2 dimensi. Kemudian fungsi aktivasi tersebut akan didapatkan *output volume* [5].

2. Pooling Layer

Lapisan ini dapat disebut juga *subsampling layer* dan diterapkan setelah *convolution layer*. Fungsi dari lapisan ini adalah untuk mengurangi parameter. Lapisan ini biasanya beroperasi menggunakan operasi *max pooling* [5].

3. Fully Connected Layer

Fully Connected Layer berfungsi untuk mengklasifikasi data masukan dengan cara mentransformasi *output* dari *convolution* dan *pooling layer* menjadi data satu dimensi. Pada dasarnya *fully connected layer* merupakan sebuah *multi layer perceptron* (MLP) yang menggunakan sebuah fungsi aktivasi untuk proses klasifikasi [5].

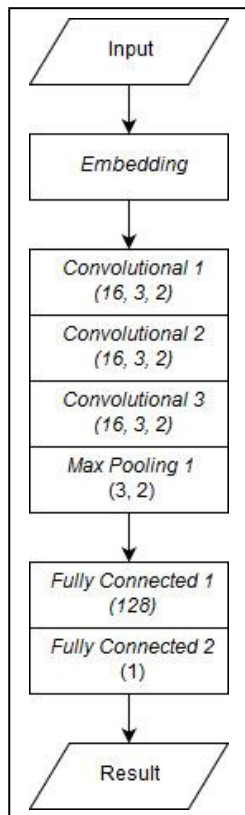
Pada NLP, *input* yang digunakan CNN adalah kalimat atau dokumen yang direpresentasikan sebagai matriks. Setiap baris matriks adalah vektor yang merepresentasikan *token* yang biasanya adalah sebuah kata. Vektor-vektor tersebut adalah hasil *word embeddings* seperti *word2vec*. Nantinya, *filter* akan dilakukan penggeseran ke seluruh baris matriks yang berisikan kata. Sehingga *width* dari *filter* biasanya sama dengan *width* dari matriks *input* namun *height* mungkin bervariasi [6].

Pada penelitian ini penulis mengusulkan arsitektur model CNN yang terdiri dari 3 *convolutional layers* dengan jumlah *filter* 16 pada tiap *layer*. Kemudian ukuran *kernel* yang digunakan adalah 3, dan *strides* 2. Sedangkan pada *maxpooling layer*, ukuran *pool* adalah 3 dan *strides* 2. Kemudian diikuti oleh 2 *fully connected layers* dengan jumlah *neuron* sebesar 128 dan 1. Arsitektur model yang diusulkan tersebut dapat dilihat pada Gambar 1.

D. Data

Data masukan atau *input* yang digunakan dalam studi kerja klasifikasi ini adalah teks dari *tweet* Twitter yang diambil menggunakan program *webcrawler*. Program *webcrawler* ini diambil dari program berbahasa Python yang sebelumnya

sudah dirancang dan dipublikasikan oleh Jefferson Henrique pada halaman GitHub miliknya.



Gambar 1. Arsitektur model CNN yang diusulkan

Terdapat 1000 data yang digunakan pada penelitian ini. Data tersebut kemudian dibagi dua kategori yaitu depresi dan normal. Penentuan kategori pada setiap data dilakukan secara manual dan mengacu pada arahan dari diskusi dengan psikolog dan beberapa teman mahasiswa jurusan psikologi. Parameter pencarian data teks dari *tweet* Twitter adalah sebagai berikut.

- *Tweet* dikirim oleh pengguna yang berlokasi di Indonesia
- *Tweet* yang telah dikirim sebelum tanggal 25 November 2018
- *Tweet* mengandung kata ataupun frasa yang telah ditentukan berdasarkan hasil diskusi oleh psikolog dan mahasiswa jurusan psikologi

Kata dan frasa kunci yang digunakan adalah sebagai berikut.

1. Bosan hidup
2. Hidup tidak berarti
3. Ingin mati
4. Kesepian
5. Lelah hidup
6. Mengakhiri hidup
7. Patah semangat
8. Putus asa
9. Saya depresi
10. Selamat tinggal dunia

Setelah data teks dari Twitter telah didapatkan, penulis melakukan pelabelan secara manual dengan membaca dan memahami maksud dari *tweet* tersebut apakah benar memiliki

indikasi depresi atau tidak. Contoh data teks yang sudah dilabeli dapat dilihat pada Tabel 1.

Data masukan kemudian diproses dan menghasilkan data keluaran berupa prediksi hasil klasifikasi depresi atau bukan depresi (normal) serta nilai akurasi dari proses klasifikasi yang telah dilakukan. Program yang dibuat juga menghasilkan *file* model CNN dari proses *training* yang telah dilakukan. Model tersebut menyimpan nilai *weight* dari proses klasifikasi sebelumnya sehingga dapat digunakan untuk melakukan klasifikasi langsung tanpa melakukan *training* terlebih dahulu.

Tabel 1.
Contoh data yang sudah dilabeli

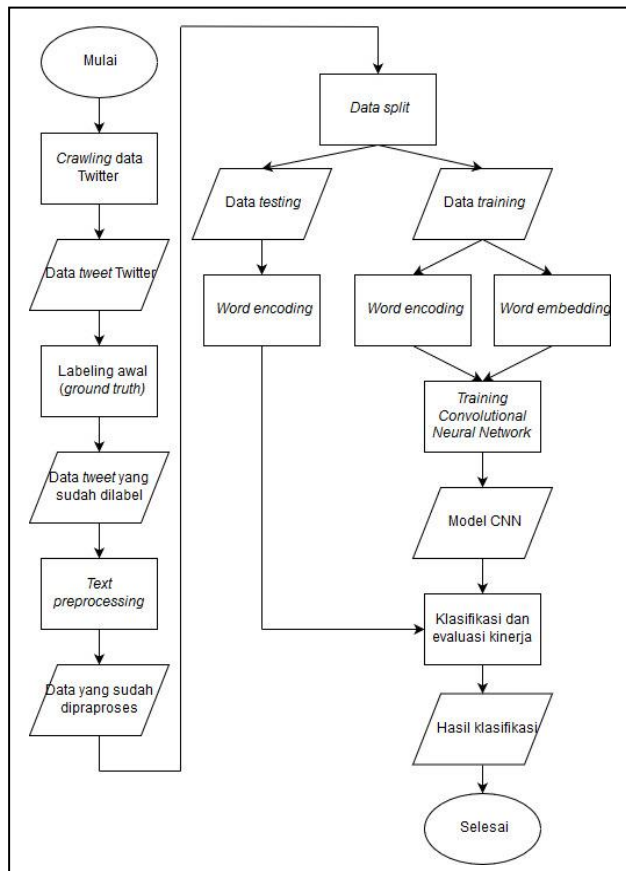
No	Data <i>tweet</i>	Kelas
1	berkutik dengan depresi akhir akhir ini sampai benar benar ingin mati. ah aku butuh teman.	Depresi
2	Ku ingin mati saja tidur selamanya daripada ku tak dapat bercerita kepada siapapun, ternyata ini menyiksa...	Depresi
3	Saat ini lah titik terendah semasa hidup saya, jiwa dan otak saya tiap hari mendidih..mempertanyakan sebenarnya siapa saya? Apa yg bisa saya lakukan? Saya depresi karena diri sendiri... pekerjaan saya amburadul, bahkan saya tidak bisa berpikir atau melakukan hal kecil...	Depresi
4	Jika kau merasa bosan dengan hidup. Mungkin kau sedang melalui titik di mana kau tidak mengharapkan apapun dalam hidupmu.	Normal
5	Hidup tidak berarti apa-apa tanpa persahabatan" - Quintus Ennius"	Normal
6	Semua orang pasti pernah merasakan jatuh dan ingin mati. Astagfirullah.	Normal

E. Metodologi

Diagram alir program secara umum dapat dilihat pada Gambar 2. Telah dihasilkan program yang dapat mendeteksi pengguna Twitter yang memiliki indikasi depresi berdasarkan *tweet* dari pengguna tersebut.

Pada percobaan ini data dibagi menjadi data *training* sebanyak 850 *tweet* dan data *testing* sebanyak 150 *tweet*. Kemudian digunakan metode *word embedding* *Word2Vec* *Skip-gram* yang di-*training* pada data *training*. Setelah itu diterapkan *word encoding* terhadap data supaya data dapat diproses oleh algoritma CNN. Kemudian dilakukan percobaan

untuk mencari arsitektur model CNN yang paling tepat untuk kasus ini. Model tersebut kemudian digunakan untuk klasifikasi data *testing* dan akurasi dibandingkan dengan hasil dari model-model CNN lainnya seperti AlexNet, ZFNet, VGG16, VGG19, dan LeNet-5.



Gambar 2. Diagram alir program secara umum

1) Text Preprocessing

Sebelum melakukan pengolahan data teks, dibutuhkan pemrosesan data terlebih dahulu untuk mengubah data teks yang tidak terstruktur jadi terstruktur. *Text preprocessing* ini dilakukan karena penyimpanan data yang terstruktur dapat membantu pengolahan data, karena dapat lebih mudah menghasilkan algoritma yang efisien.

Ada beberapa metode yang umum digunakan dalam *text preprocessing*, antara lain:

a. Tokenizing

Tokenizing adalah tahap untuk merubah teks yang panjang menjadi beberapa bagian kecil (*token*). Teks yang panjang dapat menjadi kalimat, kalimat dapat menjadi kata, dan seterusnya.

b. Cleansing

Cleansing adalah tahap untuk menghilangkan karakter-karakter yang tidak dibutuhkan untuk proses selanjutnya. *Cleansing* akan menghilangkan angka, tanda baca, dan *emoticon* pada sebuah teks.

c. Stop Word Removal

Stop word removal bertujuan untuk menghapus kata-kata yang tidak dianggap memiliki makna penting pada sebuah kalimat atau *stop word*. *Stop word* dapat berupa kata hubung maupun kata depan seperti kata “di”, “ke”, “dan”, “dengan”, serta “ini”.

d. Stemming

Stemming adalah tahap untuk menghapus imbuhan pada suatu kata dalam teks sehingga dapat menjadi kata dasar. Seperti kata “berkutik” menjadi “kutik”.

2) Word Embedding

Word embedding digunakan untuk mencari bobot dari tiap kata sebagai representasi dalam bentuk vektor dari kata tersebut. Pada percobaan akan digunakan metode *Word2Vec Skip-gram*. Metode ini mengambil vektor dari kata di sekitar kata yang sedang di-*training* untuk memahami konteks kata dari suatu kalimat [7]. Kompleksitas dari metode ini dapat dilihat pada Persamaan 1.

$$Q = C \times (D + D \times \log_2(V)) \quad (1)$$

Di mana C adalah jarak maksimal antar kata yang akan diprediksi, D adalah dimensi atau jumlah vektor representasi dari tiap kata, dan V adalah jumlah kata dalam *vocabulary*.

Pada model *Skip-gram*, *layer input* berisikan kata dari *corpus* yang akan diubah menjadi vektor representasi kata tersebut pada *layer projection*. Kemudian pada *layer output* akan dihasilkan kata-kata tetangga yang memiliki konteks sama dengan kata pada *layer input* [7]. Arsitektur model ini dapat dilihat pada Gambar 3.

3) Word Encoding

Proses ini memiliki tujuan untuk mengubah data teks menjadi data yang dapat diproses oleh model CNN. Proses ini terdiri dari dua tahapan yaitu *sequencing* dan *padding*.

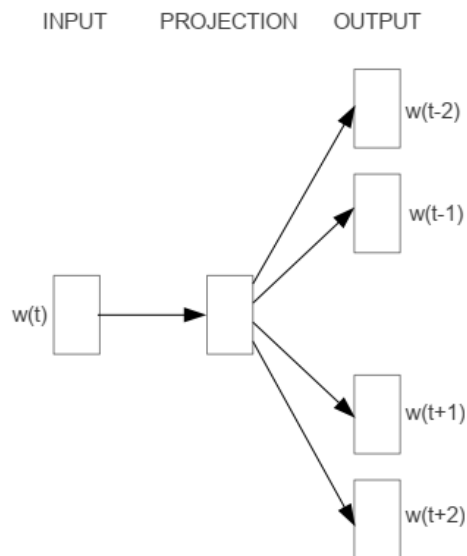
Sequencing bertujuan untuk merepresentasikan data teks menjadi angka yang menjadi indeks tiap kata pada *vocabulary*. Selanjutnya adalah proses *padding*. Proses ini akan merubah kalimat menjadi kumpulan array yang terdiri dari kumpulan kata-kata. Dalam proses ini jumlah dimensi array harus ditentukan terlebih dahulu supaya seluruh data teks yang diproses memiliki ukuran dimensi yang sama. Tahap ini penting dilakukan karena CNN membutuhkan data dengan dimensi yang tetap untuk diproses.

4) Proses Training

Percobaan menggunakan enam arsitektur model CNN yaitu AlexNet, ZFNet, VGG16, VGG19, LeNet-5, dan model yang dibangun oleh penulis. Pada tiap model CNN akan diaplikasikan *embedding layer* yang menggunakan bobot tiap kata yang sudah dihasilkan pada proses *word embedding* menjadi data masukan untuk diproses

algoritma CNN. Kemudian untuk proses validasi digunakan *Stratified 5-Fold cross validation*.

Pada proses validasi, metode *Stratified 5-Fold cross validation* membagi data *training* menjadi 5 *fold* dengan pembagian yang acak namun tiap bagian tetap memiliki representasi data secara keseluruhan. Kemudian training dilakukan pada 4 *fold* dan validasi dilakukan pada 1 *fold* sisanya. Proses ini akan berulang hingga 5 kali atau hingga semua *fold* sudah dijadikan data validasi.



Gambar 3. Arsitektur model skip-gram

III. UJI COBA

Uji coba dilakukan untuk mencari arsitektur model CNN yang tepat untuk data ini. Model terbaik yang didapat berdasarkan skenario-skenario uji coba penulis adalah model dengan 3 *layer*, *filter* berjumlah 16 pada tiap *layer*, dan menggunakan Nadam *optimizer* dengan *learning rate* sebesar 0,0001. Kemudian model tersebut dibandingkan dengan model CNN lainnya seperti AlexNet, ZFNet, VGG16, VGG19, dan LeNet-5.

Percobaan dilakukan dengan *training* pada data *training* dengan setiap model sebanyak lima kali dan diambil model terakhir dari hasil *Stratified 5-Fold cross validation*. Kemudian dilakukan klasifikasi terhadap data *testing* dan diambil rata-rata akurasi dari tiap model. Tabel 2 adalah hasil uji coba perbandingan model yang diusulkan dengan lima model CNN lainnya. Dari hasil tersebut dapat disimpulkan bahwa arsitektur yang kompleks atau memiliki banyak *layer* dan *filter* seperti AlexNet, ZFNet, VGG16, dan VGG19 cenderung *overfit* sehingga menghasilkan nilai rata-rata akurasi validasi tinggi tetapi rata-rata akurasi *testing* yang rendah.

Tabel 2.

Perbandingan model dengan yang dibangun oleh penulis dengan lima model CNN lainnya

Akurasi (%)	Alex Net	ZF Net	VGG 16	VGG 19	LeNet-5	3 Layers
Percobaan 1	59,99	59,33	61,33	60,66	70	76,66
Percobaan 2	60	59,33	61,33	65,33	69,33	75,33
Percobaan 3	57,33	63,33	63,99	64,66	70,66	75,33
Percobaan 4	60	61,99	62	62	70,66	74,66
Percobaan 5	61,33	61,33	62	60	70	74,66
Rata-rata	59,73	61,06	62,13	62,53	70,13	75,33

IV. KESIMPULAN DAN SARAN

Kesimpulan yang dapat diambil dari penelitian ini adalah sebagai berikut:

1. Arsitektur dengan *layer* yang berjumlah tiga adalah arsitektur paling optimal dengan rata-rata akurasi 72,53%.
2. Jumlah *filter* paling optimal pada arsitektur dengan *layer* berjumlah tiga adalah *filter* berjumlah 16 dengan rata-rata akurasi 72,53%.
3. *Optimizer* yang paling tepat digunakan pada arsitektur ini adalah Nadam dengan *learning rate* sebesar 0,0001 dengan rata-rata akurasi 75,33%.
4. Arsitektur yang dalam (terdiri dari banyak *layer*) dan menggunakan parameter yang banyak akan menghasilkan rata-rata akurasi yang kecil.

Beberapa saran yang dapat diberikan untuk pengembangan penelitian dengan kasus yang sama di masa yang akan datang adalah sebagai berikut :

1. Memperbanyak jumlah dan variasi data untuk menghindari *overfit* dan mendapatkan hasil akurasi yang diharapkan dari model CNN yang dibuat.
2. Metode *preprocessing* dapat diganti dengan cara lain seperti tidak melakukan *stop word removal* dan *stemming* karena metode tersebut dapat mengubah arti atau konteks data teks.
3. Dibutuhkan ketelitian yang lebih saat melakukan pelabelan data *tweet* karena tidak mudah untuk memahami maksud dari *tweet* tersebut.
4. Dapat dilakukan juga pengambilan model dengan cara berbeda dari penelitian ini, yang menggunakan model hasil *training* pada *fold* terakhir, yaitu mengambil model dengan akurasi tertinggi pada salah satu *fold*.

V. DAFTAR PUSTAKA

- [1] "Southeast Asia Digital, Social, and Mobile 2018," Asean Up, 18 April 2018. [Online]. Available: <https://aseanup.com/southeast-asia-digital-social-mobile/>. [Accessed 4 June 2018].
- [2] "Suicide Rate by Country 2018," Wolrd Population Review, [Online]. Available:

<http://www.worldpopulationreview.com/countries/suicide-rate-by-country/>. [Accessed 15 January 2019].

- [3] Y. Le Cun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and L. D. Jackel, "Handwritten Digit Recognition with a Back-Propagation Network," AT&T Laboratories, Holmdel, 1990.
- [4] A. H. Orabi, P. Buddhitha, M. H. Orabi and D. Inkpen, "Deep Learning for Depression Detection of Twitter Users," *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, pp. 88-97, 2018.
- [5] Y. Le Cun, K. Kavukcuoglu and F. Clement, "Convolutional Networks and Application in Vision," International Symposium on Circuits and Systems: Nano-Bio Circuit Fabrics and Systems, Paris, 2010.
- [6] Y. Zhang and B. C. Wallace, "A Sensitivity Analysis of (and Practitioner's Guide to) Convolutional Neural Networks for Sentence Classification," 2015.
- [7] T. Mikolov, K. Chen, G. Corrado and J. Dean, "Efficient Estimation of Word Representations in Vector Space," Mountain View, 2013.