



Klasifikasi Kepribadian dengan Metode DISC pada Twitter Menggunakan Algoritma *Artificial Neural Network*

Idris¹, Ema Utami², Anggit Dwi Hartanto³

Prodi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta

Email: idriscoklat@gmail.com¹, emma@nrar.net², anggit@amikom.ac.id³

Abstrak

Maju mundurnya suatu perusahaan biasanya didukung oleh adanya sumber daya yang handal, terutama sumber daya manusia. Perekrutan dan penempatan pegawai pada posisi yang tepat akan membawa dampak yang signifikan bagi suatu perusahaan. Di dunia ini sifat dan karakter manusia sangat beraneka ragam bentuknya. Teori DISC mengklasifikasikan kepribadian menjadi empat tipe yaitu *dominance*, *influence*, *steadiness* dan *compliance*. Perbedaan karakter setiap tipe tentu saja akan berpengaruh pada gaya perilaku, cara menghadapi tekanan hidup dan juga cara berkomunikasi baik secara langsung maupun dengan media sosial. Melalui sosial media, seseorang dapat meluapkan perasaannya melalui postingan yang diunggahnya. Dari postingan tersebut dapat dilakukan analisis mengenai karakter kepribadian yang ia dimiliki. Penelitian ini bertujuan untuk mengetahui seberapa besar akurasi analisis profiling pada Twitter sehingga bisa menjadi acuan untuk proses perekrutan pegawai. Penelitian ini menggunakan algoritma *Artificial Neural Network* untuk mengklasifikasikan 275 akun Twitter kedalam teori DISC dan mendapatkan akurasi sebesar 42,91% dari 72 skenario yang dijalankan.

Kata kunci: Kepribadian DISC, Media Sosial, Analisis Profiling, Sumber Daya Manusia

Abstract

The success or failure of a company is usually supported by the presence of reliable resources, especially human resources. Recruitment and placement of employees in the right position will have a significant impact on a company. Human nature and character are very diverse in their forms. DISC theory classifies personality into four types namely dominance, influence, steadiness, and compliance. The difference in the character of each type will of course affect the behavior style, how to deal with life pressures, and also how to communicate both directly and with social media. Through social media, a person can vent his feelings through the posts he uploaded. From these posts an analysis of the personality character he is possessed can be carried out. This study aims to find out how much the accuracy of profiling analysis on Twitter, so it can be used in the employee recruitment process. This study uses Artificial Neural Network algorithm to classify 275 Twitter accounts into DISC theory and got an accuracy of 42.91% from 72 scenarios.

Keywords: DISC, Social Media, Profiling Analysis, Human resources

A. PENDAHULUAN

Sebuah organisasi, instansi atau perusahaan akan berjalan dengan baik apabila didukung oleh sumber daya yang handal. Salah satu sumber daya yang penting adalah sumber daya manusia. Perekrutan pegawai yang tepat akan berpengaruh terhadap kinerja perusahaan tersebut. Jika kinerja meningkat, target bisa tercapai, produksi meningkat, dan penjualan pun mengalami kenaikan. Perekrutan pegawai yang tepat membutuhkan beberapa tahapan, salah satunya adalah tes kepribadian calon karyawan untuk mengetahui seberapa tinggi kecocokan antara calon karyawan dengan karakter pegawai yang dibutuhkan perusahaan. Tes kepribadian adalah sebuah tes yang dibuat untuk memberikan gambaran umum tentang gaya dan perilaku seseorang selengkap mungkin. Ada beberapa metode tes yang digunakan untuk mengetahui karakter dan kepribadian seseorang. Beberapa metode yang sering digunakan yaitu MBTI (Myers, Briggs Type Indicator), Big Five, dan DISC (Dominance, Influence, Steadiness dan Compliance). Metode DISC memberikan gambaran mengenai gaya

seseorang yang dapat memprediksi kecenderungan perilakunya di masa yang akan datang. Hal ini diperoleh dengan mengevaluasi faktor-faktor kepribadian utama yang ada dalam diri seseorang. Jika alat tes yang lengkap sering berisi ratusan pertanyaan dan membutuhkan waktu lama dalam melengkapinya, alat tes DISC hanya berupa kuisisioner berisi dua puluh empat pertanyaan, dan dapat diselesaikan dalam waktu sekitar lima belas menit. Kuisisioner tersebut dapat disajikan di atas kertas maupun dalam bentuk aplikasi sehingga penentuan nilainya dapat di otomatiskan sehingga prosesnya jauh lebih cepat. Tes DISC ini biasanya mengukur bagaimana gaya perilaku seseorang, cara berkomunikasi, cara menghadapi tekanan dan sebagainya. Orang yang mengenal dirinya sendiri dan mengerti DISC akan mudah untuk beradaptasi ke dunia orang lain yang berbeda dengan dia, gesekan-gesekan akan mengecil dan tingkat stress menurun sehingga produktivitas naik dan kesuksesan hidup juga meningkat (Shin, 2013).

Dalam 10 tahun terakhir, pengguna internet di dunia mengalami pertumbuhan yang sangat signifikan (*Digital 2020 - We Are Social*, n.d.). Karena hal tersebut, kita bisa menemui informasi tentang banyak hal di internet, terutama informasi tentang seseorang melalui laman media sosial mereka. Data yang ada di sosial media pun beragam jumlahnya, seperti status, tweet, buku yang disukai, musik yang disukai, dsb. Sehingga, kita bisa menganalisa data tersebut dan mengolahnya menjadi informasi yang lebih berguna.

Twitter adalah situs jejaring sosial mikroblogging yang digemari oleh masyarakat Indonesia. Twitter sering dijadikan objek penelitian dikarenakan sifatnya yang cepat karena dibatasinya jumlah karakter menjadi maksimal 280 karakter, sehingga tidak ada fitur “*read more*” karena semua karakter sudah tampil di layar. Oleh karena itu, informasi

yang datang menjadi sangat cepat. *Tweet* yang diposting oleh pengguna tersebut dapat di analisa dan diolah menjadi informasi yang bermanfaat dengan teknik *Profiling*.

Profiling merupakan suatu teknik yang memproses personal dan non personal data yang digunakan untuk mengembangkan pengetahuan prediktif dalam membangun suatu profil yang kemudian diterapkan dalam dasar pengambilan keputusan (Amapola et al., 2013). Data yang diambil bisa berupa data text, gambar, aktivitas pengguna, dll. Didalam penelitian ini, peneliti mengklasifikasikan kepribadian akun Twitter pengguna kedalam teori DISC menggunakan algoritma Artificial Neural Network (ANN). Beberapa penelitian yang mengkaji hubungan antara kepribadian dengan postingan di media sosial antara lain yaitu (Lhaksana, KM; Nhita, Fhira & Anggraini, 2017) melakukan penelitian dengan mengumpulkan data status facebook 100 orang pelamar dan diambil maksimal 5 status dari setiap akun. Dengan menggunakan algoritma backpropagation , pengklasifikasian kepribadian mendapatkan nilai error dengan min kesalahan 0,00001 % pada pengujian 2 dan 3. Adapun hasil pengujian yang telah dilakukan hasil perfomansi sistem dengan akurasi tertinggi 84,00 % didapatkan pada proporsi data 70:30 % data yang diujikan. Selanjutnya (Sarwani & Mahmudy, 2015) melakukan analisis terhadap tweet untuk mengetahui kepribadian seseorang yang dapat dimanfaatkan dalam perekrutan karyawan. Dengan menggunakan metode Naïve Bayess, penelitian ini menghasilkan akurasi yang sangat baik dalam melakukan klasifikasi kepribadian setelah dibandingkan dengan hasil pengklasifikasian dari pakar. Penelitian (Hutagalung, 2018) melakukan penelitian yang bertujuan untuk mengklasifikasikan jenis kepribadian melalui sosial media dengan menggunakan K-Nearest Neighbor dan teori

Big Five. Adapun hasil penelitian menunjukkan bahwa tingkat akurasi mencapai 90 %. Dari kelima kepribadian tersebut, kelas yang memiliki nilai akurasi paling tinggi adalah O, E, A, N yakni 100% sedangkan kelas yang memiliki akurasi terendah terdapat pada C yakni hanya 60%. Hal ini dikarenakan C memiliki kecenderungan yang mirip dengan kelas E.

Penelitian (Claudy et al., 2018) melakukan penelitian yang bertujuan untuk mengetahui karakter calon karyawan berdasarkan tweet yang akan diproses sehingga menghasilkan karakter kepribadian MBTI sebagai acuan untuk penempatan calon karyawan. Dari hasil penelitian tersebut diketahui bahwa algoritma K-Nearest Neighbor (KNN) dapat diimplementasikan pada sistem klasifikasi kepribadian. Total data yang digunakan yaitu 160 data yang dibagi menjadi dua dengan pembagian 50% untuk data training dan 50% untuk data uji. Hasil akhir menunjukkan bahwa, algoritma KNN menghasilkan nilai akurasi sebesar 66%. Pengklasifikasian kepribadian pengguna twitter ke dalam kelas extrovert atau introvert dengan menggunakan algoritma *Support Vector Machine* (SVM) dilakukan oleh (Fikry, 2018). Adapun hasilnya yaitu penggunaan metode SVM untuk mengklasifikasikan kepribadian pengguna twitter ke dalam extrovert dan introvert menggunakan fitur sebanyak 17 yang telah disesuaikan dengan karakteristik media sosial twitter. Setelah melakukan *grid search* dan *cross validation*, diperoleh pasangan parameter terbaik yaitu $C=1048576$ dan $\gamma=3.814697265625e-06$ pada pembagian data 80:20. Hasil pengujian dengan model terbaik tersebut memperoleh akurasi sebesar 88,89%.

Penelitian yang bertujuan untuk mengetahui hubungan antara kepribadian *Big Five* dengan pengungkapan diri secara online ini menunjukkan bahwa ada satu dimensi kepribadian *Big Five* yaitu *openness*

to experience yang mempengaruhi pengungkapan diri secara online pada pengguna media sosial. Besaran sumbangan efektif dimensi kepribadian *Big Five openness to experience* sebesar 0.193 (19.3 %) dan masih dapat mempengaruhi variabel lain yang mempengaruhi ungkapan diri selain variabel kepribadian antara lain yaitu jenis kelamin, etnis, kebangsaan, usia, efek diadik, besar kelompok topik dan penerima hubungan. Dari kelima kepribadian Big Five yang memiliki hubungan positif yang signifikan adalah *openness to experience*. Dimensi kepribadian *extraversion, agreeableness, conscientiousness dan neuroticism* tidak menunjukkan hubungan yang signifikan (Hikmah, 2017).

Pengklasifikasian jenis kepribadian seseorang ke dalam teori kepribadian Big Five berdasarkan *tweet* yang telah dibuat pengguna twitter dalam bahasa Indonesia dilakukan (Ong et al., 2017). Hasil penelitian ini menunjukkan bahwa mungkin untuk menentukan jenis kepribadian pengguna twitter tanpa literasi *tools* (LIWC, MRC), namun berdasarkan diksi yang ditulis sendiri oleh pengguna obyek penelitian ini. Penelitian saat ini membandingkan dua pengklasifikasian yang berbeda yaitu *Support Vector Machine* dan *XGBoost*. Kedua algoritma diuji dalam skenario yang berbeda yang melibatkan kemunculan minimum n-gram, skema pembobotan n-gram, penggunaan fitur LDA dan penghilangan *stopwords*. Evaluasi menggunakan 10-fold cross validations menunjukkan bahwa sistem prediksi kepribadian berdasarkan SVM berhasil mencapai akurasi rata-rata tertinggi 76.2310%, sedangkan *XGBoost* mencapai 97.9962%.

Penelitian (Damanik & Masayu Leylia Khodra, 2015) bertujuan untuk membangun model prediksi kepribadian Big Five dari pengguna twitter dengan menggunakan *Support Vector Regression* (SVR). Analisis dilakukan terhadap perilaku sosial pengguna twitter dan penggunaan

linguistic ketika menuliskan *tweet* dan biografinya untuk menentukan fitur-fitur yang paling sesuai membentuk model pembelajaran memprediksi kepribadian pengguna. Secara keseluruhan, kepribadian *Big Five* pengguna twitter dapat diprediksi dengan baik menggunakan model gabungan fitur perilaku sosial dan fitur linguistic dengan metode *open-vocabulary* bigram. Peneliti (Cahyaning, A., & Cahyono, 2015) melakukan penelitian yang bertujuan untuk mengetahui apakah terdapat perbedaan *communication privacy management* di Twitter pada remaja extrovert dan introvert. Hasil uji hipotesis dalam penelitian ini menunjukkan angka 0.01. Hal tersebut menunjukkan bahwa terdapat perbedaan antara tipe extrovert dan introvert, hasil penelitian menunjukkan bahwa tipe extrovert lebih mengungkapkan privasinya di Twitter dibandingkan dengan tipe introvert.

Klasifikasi kepribadian pengguna Facebook ke dalam kepribadian five factors dilakukan oleh penelitian (Souri et al., 2018) dengan menggunakan algoritma *Boost-Decision Tree* dan menghasilkan Akurasi sebesar 82,2%. Peneliti mendapatkan 117 partisipan dan meminta agar partisipan tersebut mengisi beberapa data dan kemudian didapatkan 100 partisipan yang valid. Kemudian peneliti menganalisa kebiasaan pengguna lewat aktivitas user yang ada di Facebooknya, seperti jumlah like, postingan, teman, dll. Peneliti menggunakan beberapa algoritma, yaitu *Naive Baiyes*, *Neural Network*, *Decision Tree*, *SVM*, *Boosting-Naïve Bayesian*, *Boosting-neural network*, *Boosting-decision tree*, dan *Boosting-SVM*. Kemudian membandingkan semua algoritma tersebut dan diketahui bahwa *Boosting-decision tree* memiliki akurasi terbaik. Selanjutnya adalah penelitian (Gómez-Adorno et al., 2016) yang terfokus pada pengembangan sebuah *lexical resource*, yang terdiri dari kamus singkatan, contractions, kata tidak baku, dan emotikon yang sering digunakan di media sosial.

Lexical resource ini nantinya akan digunakan untuk meningkatkan hasil akurasi pada penelitian-penelitian yang menggunakan algoritma *neural network* dan data yang berasal dari sosial media. *Lexical* ini dibuat karena masalah-masalah yang didapat saat memproses NLP, yaitu banyaknya bahasa gaul (bahasa yang tidak standar). Penelitian ini menggunakan Twitter dan juga kamus bahasa inggris, spanyol, belanda, dan juga italia sebagai sumber data. Hasil dari penelitian ini menunjukkan bahwa dengan menggunakan *lexical resources* ini, kualitas akurasi dari algoritma *Neural Network* bertambah ketika digunakan untuk menjalankan profiling.

Penggunaan Twitter sebagai objek penelitian dilakukan oleh penelitian (Hartanto et al., 2019). Penelitian ini bertujuan untuk memberi alternatif kepada human resources dalam memperoleh data kepribadian calon karyawan melalui akun twitter nya. Penelitian ini menggunakan algoritma *Naive Baiyes Classifier* dengan pembobotan W-IDF (*Weighted-Inverse Document Frequency*) untuk mengklasifikasikan kepribadian calon karyawan ke dalam model kepribadian DISC, serta menggunakan training data dan testing data sebanyak 120 akun. Peneliti mengacu pada *wordbag* yang telah diverifikasi psikologis dan mendapat akurasi sebesar 36.67%. Penelitian pada postingan Twitter juga dilakukan oleh penelitian (Utami et al., 2019) yang bertujuan untuk menganalisa apakah *profiling* pada twitter bisa menjadi alternatif untuk *human resource* dalam menggali kepribadian calon karyawan. Pemrosesan teks dipilih didalam penelitian ini dan dibagi menjadi empat skenario, yaitu *not stemmed-not weighted*, *stemmed-not weighted*, *not stemmed-weighted*, and *stemmed-weighted*. Hasilnya untuk masing-masing skenario adalah sebagai berikut: *not stemmed-not weighted* mendapatkan akurasi sebesar 37.41%, *stemmed-not weighted* mendapat

akurasi sebesar 30.21%, *not stemmed-weighted* mendapat akurasi sebesar 35.97%, dan *stemmed-weighted* mendapat akurasi sebesar 30.93%.

B. METODE

Di dalam *profiling analysis*, berbagai penelitian mencoba untuk menganalisa berbagai jenis data dari sosial media seseorang. Ada yang melalui status (Lhaksmana, KM; Nhita, Fhira & Anggraini, 2017), tweet (Claudy et al., 2018), aktivitas pengguna (Souri et al., 2018), dll untuk menemukan akurasi yang terbaik. Metodologi yang digunakan untuk profiling analysis dijelaskan dalam gambar alur di gambar 1.



Gambar 1. Langkah-langkah *profiling analysis*

1. Pengumpulan Data

Profiling analisis dimulai dari pengumpulan data pengguna. Pengumpulan data bisa dilakukan dengan cara menyebarkan kuisioner di internet. Kuisioner tersebut berisi pertanyaan-pertanyaan dari model yang dipilih untuk mengklasifikasikan kepribadian, contohnya adalah kuisioner yang berisi 24 pertanyaan dari model DISC. Kuisioner tersebut disajikan

dalam sebuah formulir web ataupun aplikasi web, sehingga semua orang dapat mengaksesnya. Calon partisipan akan menjawab pertanyaan-pertanyaan yang diberikan dan memasukkan username akun media sosial mereka, setelah mereka mensubmit kuisionernya, sistem akan menentukan jenis kepribadian partisipan berdasarkan pertanyaan yang sudah mereka jawab, sehingga peneliti mendapatkan data akun sosial media partisipan dan informasi jenis kepribadiannya. Setelah dirasa mendapatkan cukup partisipan, langkah selanjutnya adalah menyeleksi partisipan-partisipan tersebut agar didapat partisipan yang valid. Penyaringan tersebut berisi aturan-aturan yang dibuat berdasarkan penelitian (Liu & Zhu, 2016) yang antara lain berisi pengecekan akun media sosial partisipan apakah ada atau tidak, penghitungan total tweet atau status didalam sosial media partisipan, jika kurang dari 500 tweet, maka kurang valid. Aturan lainnya adalah sistem akan mencatat waktu yang pengguna lewati saat pengisian kuisioner, jika waktu lebih dari 10 detik maka partisipan tersebut dinyatakan valid.

Setelah mendapatkan partisipan dan data jenis kepribadian si partisipan, proses selanjutnya adalah data mining. Data mining merupakan serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data / dataset (Zulfa & Winarko, 2017). Didalam pengumpulan data terdapat berbagai cara yang dilakukan yaitu *scraping* (Hernandez-Suarez et al., 2018), *crawling* (Standard Search — Twitter Developers, n.d.), dan *streaming* (Kraft et al., 2013) (Bifet & Frank, 2010). *Scraping* adalah proses mengekstraksi informasi dari situs web secara otomatis dengan mengurai tag html dan hanya mengambil informasi yang diperlukan (Hernandez-Suarez et al., 2018). Sedangkan *crawling* dan *streaming* merupakan proses

data mining yang dilakukan melalui *Application Programming Interface* (API) suatu media sosial.

2. *Preprocessing*

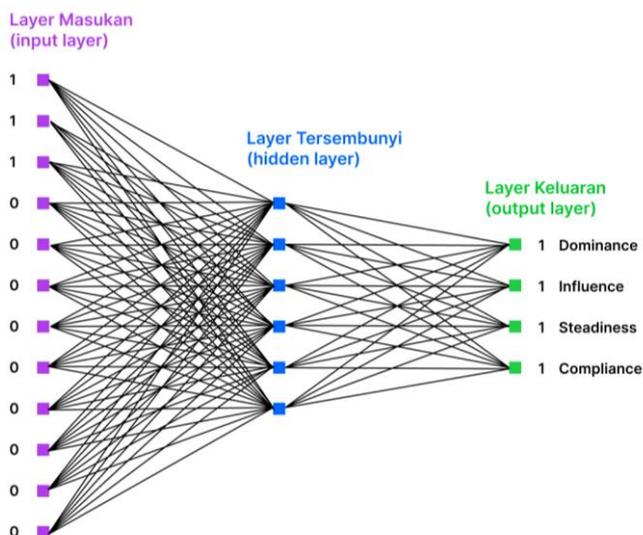
Selanjutnya adalah proses pembersihan data yang sudah dimining. Proses ini biasa juga disebut dengan *preprocessing*. *Preprocessing* dilakukan melalui beberapa langkah, yaitu: *Case folding*, Menghilangkan URL, Menghilangkan *username*, Menghilangkan tagar, Tokenisasi atau memisahkan kalimat menjadi kata-perkata, Normalisasi kata, *Stemming*, *POS Tagging* dan *POS Filtering*.

3. *Pelatihan & Pengujian*

Setelah *preprocessing*, proses berikutnya adalah *Term Frequency*. *Term Frequency* adalah proses penghitungan kata berdasarkan seberapa sering kata tersebut muncul dalam suatu dokumen. Rumus perhitungan dari term frequency adalah:

$$\text{Term Frequency (z)} = \frac{\text{Angka kemunculan kata z didalam dokumen}}{\text{Total seluruh kata didalam dokumen}}$$

Dalam penelitian ini, seluruh akun akan digabungkan menjadi satu dokumen per kelas D/I/S/C guna mencari kata yang sering muncul pada suatu kepribadian. Setelah kata yang paling sering muncul, dipilih beberapa kata yang paling sering muncul. Setelah *Term Frequency*, langkah selanjutnya adalah pembuatan model *Neural Network*. Model yang dimaksud disini adalah penentuan jumlah input layer dan hidden layer pada *Neural Network* guna mencari akurasi yang paling baik. Ilustrasi model *Neural Network* ada pada gambar 1.



Gambar 2. Model Neural Network

Model Neural Network pada umumnya terdiri dari layer masukan, layer tersembunyi, dan layer keluaran. Layer masukan dan layer tersembunyi jumlahnya harus kita sesuaikan dan optimalkan agar mendapat akurasi yang terbaik. Langkah selanjutnya adalah pembuatan dataset dengan mengkonversi data-data twitter tersebut kedalam bilangan biner, karena layer masukan *neural network* harus dalam bentuk angka, begitu juga keluarannya nanti akan berbentuk angka. Setelah akun-akun tersebut dikonversi ke biner, kita harus memvalidasi model *neural network* kita menggunakan *K-Fold Cross Validation*.

K-Fold Cross Validation adalah salah satu metode yang digunakan untuk validasi sebuah dataset dengan cara data akan di partisi atau dibagi menjadi K bagian (*fold*) segmen dengan ukuran yang sama. Proses *training* dan *testing* data dilakukan sebanyak K kali dimana pada setiap iterasi ke-*i*, segment *K_i* akan menjadi data *testing* sedangkan yang lain berperan sebagai data *training*. Penelitian ini menggunakan 10 *fold* untuk memvalidasi data.

C. HASIL DAN PEMBAHASAN

Dalam penelitian ini, kuisisioner dibuat dalam bentuk halaman website dan berisi 24 pertanyaan dan mempunyai fitur timer. Sehingga, partisipan yang menjalani tes dapat direkam waktu pengerjaannya. Kuisisioner ini kemudian disebar di Twitter dan mendapat 1.096 data partisipan yang belum difilter validitasnya. Setelah melalui serangkaian proses validitas akun berdasarkan penelitian (Liu & Zhu, 2016), jumlah partisipan yang valid adalah 275 akun. Kemudian, akun-akun tersebut dikelompokkan sesuai kepribadiannya masing-masing yaitu *Dominance*, *Influence*, *Steadiness*, maupun *Compliance*. Hasil pengelompokan dapat dilihat di tabel 1.

Tabel 1. Pengelompokan Akun Kedalam Kepribadian

Label	Jumlah Akun
D (<i>Dominance</i>)	44
I (<i>Influence</i>)	59
S (<i>Steadiness</i>)	110
C (<i>Compliance</i>)	62
Total	275

Setelah mendapatkan akun yang valid, proses selanjutnya adalah *tweet mining* menggunakan metode *scraping* (Hernandez-Suarez et al., 2018). Proses *scraping* ini mendapatkan hasil sebanyak 356.415 tweet. Hasil tersebut dapat dilihat berdasarkan kelompok pada tabel 2.

Tabel 2. Hasil Scraping

Label	Jumlah Tweet
D (<i>Dominance</i>)	53.889
I (<i>Influence</i>)	86.904
S (<i>Steadiness</i>)	147.188
C (<i>Compliance</i>)	74.230
Total	356.415

Penentuan model dilakukan dengan mengeksplorasi sebanyak mungkin skenario guna mencari *input layer*, *hidden layer*, dan *epoch* (perulangan) yang mendapat akurasi paling tinggi. Pada penelitian ini, dilakukan 72 skenario pencarian model terbaik, seluruh skenario dan hasilnya dapat dilihat di tabel 3.

Tabel 3. Skenario Pencarian Model Terbaik

No	Jumlah Input Layer	Jumlah Hidden Layer	Jumlah Epoch	Akurasi
1.	400	200	10	35.87%
2.	400	200	11	41.40%
3.	400	200	12	38.24%
4.	400	200	13	37.70%
5.	400	200	14	34.58%
6.	400	200	15	32.90%
7.	400	200	50	35.43%
8.	400	200	100	35.41%
9.	400	400	10	33.71%
10.	400	400	11	33.78%
11.	400	400	12	36.79%
12.	400	400	13	36.89%
13.	400	400	14	32.84%
14.	400	400	15	38.58%
15.	400	400	50	35.93%
16.	400	400	100	34.07%
17.	400	800	10	33.73%
18.	400	800	11	32.42%
19.	400	800	12	36.00%
20.	400	800	13	38.61%
21.	400	800	14	33.69%
22.	400	800	15	32.71%
23.	400	800	50	35.05%
24.	400	800	100	34.50%
25.	800	400	10	42.91%
26.	800	400	11	39.96%
27.	800	400	12	41.09%
28.	800	400	13	39.84%
29.	800	400	14	39.80%
30.	800	400	15	40.31%
31.	800	400	50	38.10%
32.	800	400	100	40.31%
33.	800	800	10	39.43%
34.	800	800	11	38.53%

35.	800	800	12	40.71%
36.	800	800	13	40.73%
37.	800	800	14	40.73%
38.	800	800	15	39.86%
39.	800	800	50	39.44%
40.	800	800	100	41.66%
41.	800	1600	10	38.02%
42.	800	1600	11	40.66%
43.	800	1600	12	39.37%
44.	800	1600	13	41.66%
45.	800	1600	14	40.79%
46.	800	1600	15	37.74%
47.	800	1600	50	38.04%
48.	800	1600	100	41.96%
49.	1200	600	10	28.53%
50.	1200	600	11	30.26%
51.	1200	600	12	28.57%
52.	1200	600	13	30.65%
53.	1200	600	14	30.74%
54.	1200	600	15	31.27%
55.	1200	600	50	29.40%
56.	1200	600	100	30.31%
57.	1200	1200	10	29.75%
58.	1200	1200	11	32.92%
59.	1200	1200	12	32.01%
60.	1200	1200	13	28.78%
61.	1200	1200	14	26.93%
62.	1200	1200	15	31.16%
63.	1200	1200	50	31.60%
64.	1200	1200	100	29.04%
65.	1200	2400	10	32.02%
66.	1200	2400	11	34.69%
67.	1200	2400	12	30.20%
68.	1200	2400	13	34.33%
69.	1200	2400	14	28.42%
70.	1200	2400	15	31.62%
71.	1200	2400	50	29.78%
72.	1200	2400	100	33.31%

Pada tabel 3, dapat dilihat bahwa skenario terbaik adalah skenario ke 25 dimana modelnya terdiri dari 800 *input layer*, 400 *hidden layer*, dan 10 *epoch*. Akurasi yang didapat adalah sebesar 42.91% dengan akurasi

terbesar *K-Fold Cross Validation* sebesar mencapai 62.50% dan akurasi terkecil sebesar hanya 30.43% seperti yang ditunjukkan pada gambar 3.

```
Running Fold 1 / 10  
acc: 62.50%  
Running Fold 2 / 10  
acc: 56.52%  
Running Fold 3 / 10  
acc: 30.43%  
Running Fold 4 / 10  
acc: 47.83%  
Running Fold 5 / 10  
acc: 40.91%  
Running Fold 6 / 10  
acc: 31.82%  
Running Fold 7 / 10  
acc: 40.91%  
Running Fold 8 / 10  
acc: 45.45%  
Running Fold 9 / 10  
acc: 40.91%  
Running Fold 10 / 10  
acc: 31.82%  
rata-rata akurasi adalah 42.91%
```

Gambar 3. Hasil *K-Fold Cross Validation*

D. PENUTUP

Simpulan dan Saran

Penelitian ini membahas tentang pengklasifikasian akun Twitter kedalam teori kepribadian DISC menggunakan algoritma *Artificial Neural Network* dengan metode *Backpropagation*. Pengklasifikasian akun dilakukan dengan menganalisa teks yang berasal dari *tweet* partisipan. Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma *Artificial Neural Network* dapat menghasilkan akurasi terbaik sebesar 42.91%. Penelitian selanjutnya diharapkan menggunakan *variable-variable* lain selain teks, seperti emoji, gambar, tagar, lokasi, dan lain-lain agar dapat meningkatkan akurasi.

DAFTAR PUSTAKA

- Amapola, F., Bosco, F., Cafiero, G., Angelo, E. D., & Unicri, Y. S. (2013). Defining Profiling Working Paper. *SSRN ELibrary*, 1–40.
- Bifet, A., & Frank, E. (2010). Sentiment knowledge discovery in Twitter streaming data. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-642-16184-1_1.
- Cahyaning, A., & Cahyono, R. (2015). Perbedaan Communication Privacy Management di Media Sosial Twitter pada Remaja dengan Tipe Kepribadian Extravert dan Introvert. *Kesehatan Lingkungan*.
- Claudy, Y. I., Perdana, R. S., & Fauzi, M. A. (2018). Klasifikasi Dokumen Twitter Untuk Mengetahui Karakter Calon Karyawan Menggunakan Algoritme K-Nearest Neighbor (KNN). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(8), 2761–2765. <https://www.researchgate.net/publication/322959490>.
- Damanik, A. T., & Masayu Leylia Khodra. (2015). Prediksi Kepribadian Big 5 Pengguna Twitter dengan Support Vector Regression. *Jurnal Cybermatika*.
- Digital 2020 - We Are Social*. (n.d.). Retrieved May 17, 2020, from <https://wearesocial.com/digital-2020>.
- Fikry, M. (2018). Ekstrover atau Introver : Klasifikasi Kepribadian Pengguna Twitter dengan Menggunakan Metode Support Vector Machine. *Jurnal Sains Dan Teknologi Industri*, 16(1), 72. <https://doi.org/10.24014/sitekin.v16i1.5326>.
- Gómez-Adorno, H., Markov, I., Sidorov, G., Posadas-Durán, J.-P., Sanchez-Perez, M. A., & Chanona-Hernandez, L. (2016). Improving Feature Representation Based on a Neural Network for Author Profiling in Social Media Texts. *Computational Intelligence and Neuroscience*, 2016, 1638936. <https://doi.org/10.1155/2016/1638936>.

- Hartanto, A. D., Utami, E., Adi, S., & Hudnanto, H. S. (2019). Job seeker profile classification of twitter data using the naïve bayes classifier algorithm based on the DISC method. *2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering, ICITISEE 2019*. <https://doi.org/10.1109/ICITISEE48480.2019.9003963>.
- Hernandez-Suarez, A., Sanchez-Perez, G., Toscano-Medina, K., Martinez-Hernandez, V., Sanchez, V., & Perez-Meana, H. (2018). *A Web Scraping Methodology for Bypassing Twitter API Restrictions*. 1–7. <http://arxiv.org/abs/1803.09875>.
- Hikmah, N. M. (2017). *Hubungan Kepribadian Big Five Dengan Pengungkapan Diri Secara Online Pengguna Media Sosial*.
- Hutagalung, N. (2018). *Klasifikasi Tipe Kepribadian Pengguna Sosial Media Berdasarkan Teori Big Five Menggunakan K-Nearest Neighbor*. 14–15.
- Kraft, T., Wang, D. X., Delawder, J., Dou, W., Yu, L., & Ribarsky, W. (2013). Less After-the-Fact: Investigative visual analysis of events from streaming twitter. *IEEE Symposium on Large Data Analysis and Visualization 2013, LDAV 2013 - Proceedings*. <https://doi.org/10.1109/LDAV.2013.6675163>.
- Lhaksmana, KM; Nhita, Fhira & Anggraini, D. (2017). Klasifikasi Kepribadian Berdasarkan Status Facebook Menggunakan Metode Backpropagation. *E-Proceeding of Engineering*, 4(3), 5174–5183.
- Liu, X., & Zhu, T. (2016). Deep learning for constructing microblog behavior representation to identify social media user's personality. *PeerJ Computer Science*. <https://doi.org/10.7717/peerj-cs.81>.
- Ong, V., Rahmanto, A. D. S., Williem, W., Suhartono, D., Nugroho, A. E., Andangsari, E. W., & Suprayogi, M. N. (2017). Personality prediction based on Twitter information in Bahasa Indonesia. *Proceedings of the 2017 Federated Conference on Computer Science and Information Systems, FedCSIS 2017*. <https://doi.org/10.15439/2017F359>.

- Sarwani, M. Z., & Mahmudy, W. F. (2015). Analisa Twitter Untuk Mengetahui Karakter. *Seminar Nasional Sistem Informasi Indonesia, November, 2–3*.
- Shin, E. (2013). *The DISC Codes: Cara Cepat Menguasai Kode Sukses Manusia*. Bor Monda Tour.
- Souri, A., Hosseinpour, S., & Rahmani, A. M. (2018). Personality classification based on profiles of social networks' users and the five-factor model of personality. *Human-Centric Computing and Information Sciences*, 8(1). <https://doi.org/10.1186/s13673-018-0147-4>.
- Standard search — Twitter Developers*. (n.d.). Retrieved May 17, 2020, from <https://developer.twitter.com/en/docs/tweets/search/overview/standard>.
- Utami, E., Hartanto, A. D., Adi, S., Oyong, I., & Raharjo, S. (2019). Profiling analysis of DISC personality traits based on Twitter posts in Bahasa Indonesia. *Journal of King Saud University - Computer and Information Sciences*. <https://doi.org/10.1016/j.jksuci.2019.10.008>.
- Zulfa, I., & Winarko, E. (2017). Sentimen Analisis Tweet Berbahasa Indonesia Dengan Deep Belief Network. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*. <https://doi.org/10.22146/ijccs.24716>.

