

ANALISIS KARAKTER KEPERIBADIAN SESEORANG MENGGUNAKAN

ALGORITMA *NAÏVE BAYESS CLASSIFIER*

(Studi Kasus Twitter Resmi Basuki Tjahaja Purnama dan Anies Baswedan)

SKRIPSI

Untuk memenuhi sebagian persyaratan untuk mencapai derajat S-1

Program Studi Teknik Informatika



Disusun Oleh:

Ire Icca Agustiorini Harsehanto

14650007

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA

YOGYAKARTA

2018



PENGESAHAN SKRIPSI/TUGAS AKHIR

Nomor : B-1189/UIN.02/D.ST/PP.01.1/08/2018

Skripsi/Tugas Akhir dengan judul : "Analisis Karakter Kepribadian Seseorang Menggunakan Algoritma *Naive Bayess Classifier* (Studi Kasus Twitter Resmi Basuki Tjahaja Purnama dan Anies Baswedan)"

Yang dipersiapkan dan disusun oleh :
Nama : Ire Icca Agustiorini Harsehanto
NIM : 14650007
Telah dimunaqasyahkan pada : 10 Agustus 2018
Nilai Munaqasyah : A/B
Dan dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga

TIM MUNAQASYAH :

Ketua Sidang

M. Didik R Wahyudi, M.T
NIP. 19760812 200901 1 015

Penguji I

Agung Fatwanto, Ph.D
NIP.19770103 200501 1 003

Penguji II

Nurochman, S.Kom.M.Kom.
NIP.19801223 200901 1 007

Yogyakarta, 2016
UIN Sunan Kalijaga
Fakultas Sains dan Teknologi



Dr. Murtono, M.Si
NIP. 19691212 200003 1 001



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi

Lamp :-

Kepada
Yth. Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga Yogyakarta
di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Ire Icca Agustiorini Harsehanto
NIM : 14650007
Judul Skripsi : "Analisis Karakter Kepribadian Seseorang Menggunakan Algoritma *Naive Bayess Classifier* (Studi Kasus Twitter Resmi Bapak Basuki & Bapak Anies)"

sudah dapat diajukan kembali kepada Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Teknik Informatika.

Dengan ini kami berharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqsyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 08 Juli 2018
Pembimbing

M. Didik Rohmad Wahyudi, S.T., MT.
NIP. 19760812 200901 1 015

SURAT PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan dibawah ini:

Nama : Ire Icca Agustiorini Harsehanto
NIM : 14650007
Program Studi : Teknik Informatika
Fakultas : Sains dan Teknologi

Dengan ini saya menyatakan bahwa skripsi saya yang berjudul “**Analisis Karakter Kepribadian Seseorang Menggunakan Algoritma Naïve Bayess Classifier (Studi Kasus Twitter Resmi Basuki Tjahaja Purnama dan Anies Baswedan)**” merupakan hasil penelitian saya sendiri, tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi, dan bukan plagiasi karya orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 08 Juli 2018

Yang menyatakan,



Ire Icca Agustiorini Harsehanto
NIM. 14650007

KATA PENGANTAR

Alhamdulillah *abilalamin*, puji syukur kehadiran Allah SWT yang telah melimpahkan rahmat, hidayah serta karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “**Analisis Karakter Kepribadian Seseorang Menggunakan Algoritma *Naïve Bayess Classifier* (Studi Kasus Twitter Resmi Basuki Tjahaja Purnama dan Anies Baswedan)**”. Penulis menyadari bahwa penyusunan laporan penelitian ini tidak akan berjalan lancar tanpa dukungan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terimakasih banyak kepada:

1. Bapak Prof. Yudian Wahyudi, MA, Ph.D, selaku Rektor UIN Sunan Kalijaga Yogyakarta.
2. Bapak Dr. Murtono, M.Si selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga.
3. Bapak Dr. Bambang Sugiantoro, MT. selaku Ketua Program Studi Teknik Informatika UIN Sunan Kalijaga Yogyakarta sekaligus pembimbing akademik.
4. Bapak Muhammad Didik Rohmad Wahyudi, S.T., MT. selaku pembimbing tugas akhir, yang telah memberikan banyak sekali bantuan, arahan, saran, nasihat, waktu serta masukan kepada penulis dalam menyusun skripsi ini.
5. Bapak Sumarsono, S.T, M.Kom, selaku Dosen Pembimbing Akademik selama masa perkuliahan.
6. Seluruh dosen Program Studi Teknik Informatika UIN Sunan Kalijaga Yogyakarta yang saya hormati yang telah memberikan banyak ilmunya,

tidak hanya ilmu pengetahuan tapi juga pengalaman dan nasihat yang sangat berguna, kepada saya selama perkuliahan..

7. Kedua orangtua, Mamas-adik dan keluarga yang memberikan dukungan, semangat, nasihat-nasihat, pengalaman, dan kebahagiaan. Sehingga membuat saya semangat dalam mengerjakan skripsi ini.
8. Seluruh teman-teman seperjuangan saya, Mahasiswa Teknik Informatika 2014 yang telah menemani saya berproses selama ini, dan telah memberikan banyak sekali pengalaman, motivasi dan pelajaran berharga pada saya.
9. Serta kepada semua pihak yang tidak bisa saya sebutkan satu persatu, yang sudah membantu keberlangsungan skripsi ini secara langsung maupun tidak langsung.

Semoga Allah SWT senantiasa memberikan pahala yang setimpal terhadap semua dukungan, semangat yang telah diberikan hingga terselesaikannya Tugas Akhir ini. Penulis menyadari masih banyak sekali kekurangan dan kelemahan dalam skripsi ini. Oleh karena itu, segala kritik dan saran senantiasa penulis harapkan dari pembaca. Akhir kata, semoga skripsi ini dapat bermanfaat bagi dunia pendidikan dan segala pihak, terutama bagi penulis sendiri.

Yogyakarta, 10 Agustus 2018
Penyusun,

Ireicca Agustiorini H.

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur yang teramat besar dan teramat dalam, skripsi ini saya persembahkan untuk:

1. Kepada kedua orang tua tercinta, Bapak Suharto dan Ibu Mintarsih yang telah mendidik dan membesarkan saya. Terimakasih atas doa dan dukungannya serta kasih sayang yang tiada terkira, sehingga saya dapat menyelesaikan skripsi ini.
2. Kedua saudara laki - laki saya Mas Sandy dan Adek Aldi. Terimakasih telah memberikan dukungan secara tidak langsung termasuk dukungan dalam bentuk semangat.
3. Sahabat saya yang jauh disana Nilam Sari dan Sepupu saya Nitenia Nishi Hanazqa. Terimakasih selalu memberikan dukungan semangat dan lainnya.
4. Ridwan Wulida Siam yang telah memberikan masukan, saran, dan dukungan semangat. Terimakasih atas bantuannya dalam proses penyelesaian skripsi ini.
5. Teman seperjuangan skripsi saya Lusiana Lestari. Terimakasih untuk masukan, saran, bantuan serta dukungan semangatnya selama ini.
6. Ciwi-ciwi Informatika 2014, yang selalu ribut di group. Terimakasih atas bantuan dan dukungannya.
7. Keluarga besar Teknik Informatika angkatan 2014.

HALAMAN MOTTO

“Dan hanya kepada Allah hendaknya kamu berharap”

(QS. Al-Insyirah 94:8)

“Berpikirlah seperti seorang ratu. Seorang ratu tidak akan takut untuk gagal.
Kegagalan adalah sebuah batu loncatan menuju keberhasilan”

-Oprah Winfrey-



DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
SURAT PERSETUJUN SKIRPSI/TUGAS AKHIR.....	iii
PERNYATAAN KEASLIAN SKRIPSI	iv
KATA PENGANTAR	v
HALAMAN PERSEMBAHAN	vii
HALAMAN MOTTO.....	viii
DAFTAR ISI	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xiv
INTISARI	xv
ABSTRACT	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.2 Tujuan Penelitian.....	3
1.3 Batasan Masalah.....	3
1.5 Manfaat Penelitian.....	3
1.6 Keaslian Penelitian	4
1.7 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	6
2.1 Tinjauan Pustaka	6
2.2 Landasan Teori	9
2.2.1 Kepribadian	9
2.2.2 Teori Kepribadian	10
2.2.3 Test Psikologi.....	15
2.2.4 Media Sosial.....	16

2.2.5	Twitter	16
2.2.6	Web Crawler	17
2.2.7	PHP	18
2.2.8	Text Mining.....	19
2.2.9	Classifier	20
2.2.10	Supervised Learning.....	21
2.2.11	Naïve Bayess Classifier.....	21
BAB III METODE PENELITIAN		26
3.1	Studi Literatur.....	27
3.2	Persiapan Data.....	27
3.3	Seleksi Data.....	27
3.4	Analisis dan Perancangan.....	28
3.5	Implementasi	28
BAB IV HASIL DAN PEMBAHASAN.....		30
4.1	Analisis Kebutuhan	30
4.2	Pengambilan Data.....	31
4.3	Seleksi dan Pelabelan Data.....	32
4.3.1	Seleksi Data.....	32
4.3.2	Pelabelan Data.....	33
4.4	Pre-Processing	35
4.4.1	Cleansing.....	35
4.4.2	Case Folding	36
4.4.3	Tokenisasi	38
4.4.4	Slang Words	39
4.4.5	Stop-Words Removal.....	41
4.4.6	Stemming	43
4.5	Proses Analisis Karakter Kepribadian.....	44
4.5.1	Pembobotan Kata Dengan Term Frequency	44
4.5.2	Proses Naïve Bayes Classifier.....	52
4.6	Hasil.....	54
BAB V PENUTUP		56

5.1	Kesimpulan.....	56
5.2	Saran.....	56
	DAFTAR PUSTAKA.....	58
	LAMPIRAN	60
	LAMPIRAN I : Tabel Daftar Slang-Words	60
	LAMPIRAN II : Tabel Daftar <i>StopWords-Removal</i>	62



DAFTAR GAMBAR

Gambar 2.1 Crawling Data Twitter	26
Gambar 3.1 Alur Metode Penelitian.....	26
Gambar 4.1 Alur Proses Penelitian	30
Gambar 4.2 Flowchart Proses Naïve Bayess Classifier	53
Gambar 4.2 Diagram Hasil Klasifikasi @basuki_btp	54
Gambar 4.2 Diagram Hasil Klasifikasi @aniesbaswedan.....	55



DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu.....	8
Tabel 4.1 Contoh Data Tweet.....	31
Tabel 4.2 Pelabelan Manual	33
Tabel 4.3 Hasil <i>Cleansing</i> Data Tweet.....	35
Tabel 4.4 Hasil <i>Case Folding</i> Data Tweet	37
Tabel 4.5 Hasil <i>Tokenisasi</i> Data Tweet.....	38
Tabel 4.6 Hasil <i>Slang World</i> Data Tweet.....	40
Tabel 4.7 Hasil <i>Stopworld Removal</i> Data Tweet	42
Tabel 4.8 Hasil <i>Stemming</i> Data Tweet	43
Tabel 4.9 Data Tweet	44
Tabel 4.10 <i>Term Frequency</i>	45
Tabel 4.11 Nilai Peluang <i>Term</i> Pada Kelas <i>Openness</i>	48
Tabel 4.12 Nilai Peluang <i>Term</i> Pada Kelas <i>Neuroticism</i>	50

DAFTAR LAMPIRAN

LAMPIRAN I : Tabel Daftar *Slang-Words*.....**Error! Bookmark not defined.**

LAMPIRAN II : Tabel Daftar *StopWords-Removal* 62



**ANALISIS KARAKTER KEPRIBADIAN SESEORANG MENGGUNAKAN
ALGORITMA *NAÏVE BAYESS CLASSIFIER* (STUDI KASUS TWITTER
RESMI BASUKI TJAHAJA PURNAMA DAN ANIES BASWEDAN)**

IRE ICCA AGUSTIORINI HARSEHANTO

14650007

INTISARI

Ada banyak cara yang dapat dilakukan untuk mengetahui bagaimana karakter kepribadian seseorang. Salah satu caranya yaitu, dengan melakukan tes psikologi atau wawancara. Zaman sekarang media sosial dapat memberikan informasi tentang karakter kepribadian seseorang yang dapat dilihat dari gaya postingan yang dibagikan.

Penelitian ini menggunakan data dari media sosial Twitter berdasarkan hasil tweet dari *user_timeline* @basuki_btp dan @aniesbaswedan. Penelitian ini menggunakan 2100 data tweet. Data yang sudah terkumpul kemudian dilakukan *Pre-Processing* terlebih dahulu serta diberi label secara manual. Proses selanjutnya yaitu pengklasifikasian menggunakan Algoritma *Naïve Bayess Classifier* dengan menggunakan Teori Kepribadian Lima Besar.

Berdasarkan hasil pengujian dengan menggunakan 500 data tweet sebagai data *training* dan 1600 data tweet sebagai data *testing*. Hasil klasifikasi yang diperoleh dengan menggunakan Metode *Naïve Bayes Classifier* dan dikelompokkan dalam "Lima Besar" kelompok kepribadian, yaitu: *Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness* dan *Neuroticism* pada data tweet dalam bahasa Indonesia.

Kata Kunci: *Naïve Bayess Classifier*, Media Sosial, *Twitter*, Analisis kepribadian, Teori Kepribadian, Kepribadian Lima Besar

**ANALYSIS CHARACTER OF PERSONALITY USING NAÏVE
BAYESS CLASSIFIER ALGORITHM (CASE STUDY OF OFFICIAL
TWITTER BASUKI TJAHAJA PURNAMA AND ANIES
BASWEDAN)**

IRE ICCA AGUSTIORINI HARSEHANTO

14650007

ABSTRACT

There are many ways that can be done to find out how the character of a person's personality. One way is to do a psychological test or interview. Nowadays social media can provide information about the character of a person's personality that can be seen from the style of the posts shared.

This research uses data from social media Twitter based on the results of tweets from user_timeline @basuki_btp and @aniesbaswedan. This study uses 2100 tweet data. Data that has been collected is then pre-processed first and labeled manually. The next process is classification using the Naïve Bayess Classifier Algorithm using the Big Five Personality Theory.

Based on the test results using 500 tweet data as training data and 1600 tweet data as testing data. The classification results obtained by using the Naïve Bayes Classifier Method and grouped in the "Big Five" personality groups: Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism on tweet data in Indonesian.

Keywords: Naïve Bayess Classifier, Social Media, Twitter, Personality Analysis, Personality Theory, Big Five Personality

BAB I

PENDAHULUAN

1.1 Latar Belakang

Salah satu cara untuk mengetahui kepribadian seseorang yaitu dengan melakukan tes psikologi. Tes psikologi yang dilakukan saat ini kebanyakan melalui tes tertulis ataupun tes wawancara. Saat ini banyak penelitian tentang kepribadian seseorang yang dilakukan menggunakan media sosial, karena media sosial menjadi suatu kebutuhan yang sangat penting bagi masyarakat, terutama masyarakat modern masa kini yang selalu membagikan kegiatan sehari-harinya di media sosial.

Media sosial adalah sebuah media untuk bersosialisasi satu sama lain dan dilakukan secara *online* yang memungkinkan manusia untuk saling berinteraksi tanpa dibatasi ruang dan waktu. Media sosial pada saat ini sangat mempermudah penggunaannya dalam berbagai hal. Tidak hanya untuk memberi kabar penting saja, akan tetapi media sosial dapat digunakan untuk membagikan sebuah pesan singkat, mengirim pesan, berkomentar, menjalin pertemanan, berkirim foto, ruang untuk saling bertukar pendapat dan lain sebagainya. Terbukti hanya dalam beberapa tahun Twitter telah menjadi salah satu media sosial populer dikalangan masyarakat sampai saat ini.

Twitter juga sering digunakan sebagai media untuk mempublikasikan kegiatan sehari-hari melalui sebuah pesan singkat yang terdiri dari maksimal 280 karakter (disebut tweet). Tweet sendiri bisa terdiri dari pesan teks dan foto. Melalui tweet inilah pengguna Twitter dapat berinteraksi lebih dekat

dengan pengguna lainnya dengan mengirimkan tentang apa yang sedang mereka pikirkan, apa yang sedang dilakukan, tentang kejadian yang baru saja terjadi, tentang berita terkini serta hal lainnya sehingga banyak pengguna Twitter yang secara tidak sadar memberikan informasi tentang kepribadiannya melalui tweet yang mereka buat dengan bahasa yang alami.

Penelitian ini menggunakan Twitter sebagai media untuk menganalisis kepribadian seseorang. Dalam proses menganalisis kepribadian seseorang melalui Twitter dibutuhkan metodologi yang tepat untuk mendapatkan hasil yang akurat. Tweet merupakan kumpulan kata yang tidak baku sehingga dibutuhkan perlakuan khusus untuk mendapatkan data yang bisa diproses. Oleh sebab itu, dalam pengolahan data diperlukan proses *pre-processing* terlebih dahulu yang kemudian bisa dilakukan pengklasifikasian data. Dalam penelitian ini, metode yang digunakan adalah *Naïve Bayes Classifier* dan dikelompokkan dalam "Lima Besar" kelompok kepribadian. Metode tersebut dipilih karena sederhana dan memberikan kemudahan dalam proses pengolahan data serta memberikan tingkat akurasi yang cukup akurat.

1.2 Rumusan Masalah

Berdasarkan penjelasan latar belakang diatas, maka rumusan masalah yang akan dibahas adalah, Bagaimana melakukan analisis terhadap data tweet pengguna untuk mengetahui karakter kepribadian seseorang menggunakan algoritma *Naïve Bayess Classifier* dan dikelompokkan dalam "Lima Besar" Kelompok Kepribadian ?

1.2 Tujuan Penelitian

Tujuan yang ingin dicapai dari penelitian ini adalah, melakukan analisis berdasarkan data tweet pengguna untuk mengetahui karakter kepribadian seseorang menggunakan algoritma *Naïve Bayess Classifier* dan dikelompokkan dalam "Lima Besar" kelompok kepribadian.

1.3 Batasan Masalah

Agar penelitian lebih terarah dan tidak menyimpang dari rumusan masalah yang ada, maka batasan masalah dari penelitian ini adalah:

1. Metode klasifikasi yang digunakan adalah *Naïve Bayess Classifier*.
2. Data yang dianalisis hanya tweet asli pengguna (*user timeline*).
3. Waktu pengambilan data dilakukan secara random, dengan rentan waktu 24 jam.
4. Hasil penelitian merupakan hasil analisis untuk mengetahui karakter kepribadian seseorang pada media sosial Twitter.
5. Klasifikasi menggunakan teori kepribadian (*Big Five Personality*)
6. Bahasa pemrograman yang digunakan yaitu PHP.

1.5 Manfaat Penelitian

Dengan tercapainya tujuan penelitian diatas, maka diharapkan akan bermanfaat bagi semua pihak, manfaat penelitian ini adalah:

1. Dengan menggunakan algoritma *Naïve Bayess Classifier* dapat mengkasifikasikan karakter seseorang melalui Tweet-nya.

2. Bagi UIN Sunan Kalijaga, dapat digunakan sebagai tambahan referensi yang terkait penelitian dengan tema '*Text Mining*'.
3. Bagi Penulis, dapat memberikan pengetahuan yang baru, bahwa karakteristik kepribadian seseorang dapat dilihat dari status media sosialnya.
4. Bagi Pembaca, dengan adanya penulisan ini dapat bermanfaat sebagai sumber pembelajaran untuk penelitian selanjutnya serta untuk menambah wawasan.

1.6 Keaslian Penelitian

Penelitian yang berkaitan tentang analisis dan klasifikasi sentiment atau karakter kepribadian seseorang pada media sosial Twitter, hingga saat ini sudah banyak dilakukan oleh peneliti sebelumnya. Namun, berdasarkan referensi dan tinjauan pustaka, penelitian yang diajukan sebagai Tugas Akhir S1 pada program studi Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga mengenai. "Analisis Karakter Kepribadian Seseorang Menggunakan Algoritma *Naïve Bayess Classifier* (Studi Kasus Twitter Resmi Basuki Tjahaja Purnama dan Anies Baswedan)", belum pernah dilakukan di UIN Sunan Kalijaga Yogyakarta.

1.7 Sistematika Penulisan

Untuk memberikan gambaran dan kerangka yang jelas mengenai pokok bahasan setiap bab dalam penelitian ini, maka diperlukan sistematika

penulisan. Berikut ini gambaran sistematika penulisan pada masing-masing bab:

BAB I : PENDAHULUAN

Bab ini berisikan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, keaslian penelitian dan sistematika penulisan skripsi.

BAB II : TINJUAN PUSTAKA DAN LANDASAN TEORI

Bab ini menjelaskan tentang tinjauan pustaka dan landasan teori yang berhubungan dengan topik yang akan dibahas dalam penelitian tugas akhir.

BAB III : METODE PENELITIAN

Bab ini membahas tentang metode penelitian yang digunakan dan langkah-langkah yang harus dilakukan untuk mencapai tujuan dalam menyelesaikan penelitian tugas akhir.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini membahas hasil penelitian yang dicapai dari proses penyeleksian data, pengolahan data sampai hasil akhir yang diperoleh.

BAB V : PENUTUP

Bab ini membahas kesimpulan yang dihasilkan serta saran yang akan diberikan berdasarkan hasil yang telah dicapai, sehingga dapat digunakan sebagai bahan pertimbangan bagi pihak-pihak yang berkepentingan, serta kemungkinan perkembangan untuk penelitian selanjutnya.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan menggunakan metode *Naïve Bayes Classifier* dan dikelompokkan dalam "Lima Besar" Kelompok Kepribadian (*Openness, Conscientiousness, Extraversion, Agreeableness* dan *Neuroticism*) pada data tweet dalam bahasa Indonesia diperoleh kesimpulan sebagai berikut :

1. Basuki Tjahaja purnama memperoleh nilai *trait* yang paling besar atau dominan pada *Agreeableness* yaitu 26.8%, kemudian *Conscientiousness* 22.1%, *Neuroticism* 20.3%, *Extraversion* 18.4% dan untuk *Openness* 12.5%.
2. Anies Baswedan memperoleh nilai *trait* yang paling besar atau dominan pada *Agreeableness* yaitu 25.8%, kemudian *Extraversion* 23.9%, *Neuroticism* 20.6%, *Conscientiousness* 17.1% dan untuk *Openness* 12.6%.

Hasilnya menunjukkan *trait* yang paling besar atau dominan sama sama berada di *Agreeableness* dan *trait* terkecilnya pada *Openness*.

5.2 Saran

Pada penelitian ini tentunya tidak terlepas dari kekurangan dan kelemahan. Oleh karena itu, peneliti menyarankan adanya perbaikan-perbaikan :

1. Sistem diharapkan dapat menganalisis tidak hanya bahasa Indonesia saja, tetapi dapat menganalisis bahasa daerah, dan bahasa asing seperti bahasa Inggris, bahasa Arab, dan bahasa lainnya.
2. Pada penelitian berikutnya diharapkan menggunakan dua metode klasifikasi atau lebih sebagai pembandingan.



DAFTAR PUSTAKA

- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Barrick, M.R. & Mount, M.K. (1991). The Big Five Personality Dimensions and Job Performance: A Meta-Analysis. *Personnel Psychology*, 44 (1), 1-27.
- Even, Y., and Zohar. 2002. *Introduction to Text Mining*. Automated Learning Group National Center For Supercomputing Applicarions, University of Illionis.
- Feist, J. & . Feist, G.J. (2010). *Theories of personality*. Jakarta: Salemba Humanika.
- GA Buntoro. (2017). *Analisis Sentiment Calon Gubernur DKI Jakrta 2017 di Twitter*. Teknik Informatika, Universitas Muhammadiyah Ponorogo
- Han, Jiawei, Micheline Kamber, and Jian Pei. 2012. "Data Mining : Concepts and Techniques Third Edition." Waltham, MA: MorganKaufmann.<http://studentnationaleducation.blogspot.com/2012/1/1/penjelasan-tentangtoefl-materi-toefl.html>.
- Hidayatullah, Ahmad Fathan. 2014. *Analisis Sentimen Dan Klasifikasi Klasifikasi Kategori Terhadap Tokoh Publik Pada Data Twitter Menggunakan Naive Bayes Classifier*. MCs Thesis, Yogyakarta: Universitas Gajah Mada.
- K. Haewoon, L. Chang Hyun, P. Hosung and M. Sue. "What is Twitter, a Social network or a News Media?". International Conference WWW 2010. Raleigh, North California, USA. ACM 978-1-60558-799-8/10/04.
- Liu, Bing. 2012. *Sentiment Analysis and Opini Mining*. Morgan & Claypool Publisher.
- Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schutze. 2008. *Introduction to Information Retrieval*. United Kingdom: Cambridge University Press.
- McCrae, R. R. (2011). 5 Years of Progress: A Reply to Block. *Journal of Research in Personality* , 108–113.

- Mustafa, M. S., Ramadhan, M. R., & Thenata, A. P. (2017). Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Citec Journal*, 151-162.
- M. Zoqi, Wayan. (2015). *Analisis Twitter untuk mengetahui karakter seseorang menggunakan algoritma naïve bayess classifier*. Teknologi Informasi dan Ilmu Komputer, Universitas Brawijaya
- Natalius, S. 2010. "Metode Naive Bayes Classifier dan Penggunaannya pada Klasifikasi Dokumen." *Makalah II2092 Probabilitas dan Statistik - Sem. ITahun 2010/2011*.
- Pervin, L. A., Cervone, D., & Oliver, P. J. (2012). Psikologi kepribadian: Teori dan Penelitian (edisi ke-9). Kencana Prenada Media Group. Jakarta.
- Saraswati, Ni Wayan Sumartini. 2011. *Text Mining Dengan Metode Naive Bayes Classifier dan Support Vector Mechine Untuk Sentiment Analysis*. Thesis, Denpasar: Universitas Udayana.
- Suyanto. (2017). Data Mining untuk Klasifikasi dan Klasterisasi Data. Bandung: Informatika Bandung.
- V. George, S Antonia and G. Dimitros. 2014 "A Faceted Crawler for the Twitter Service". Thessaloniki, Greece.
- Widodo, P. P., Handayanto, R. T., & Herlawati. (2013). Penerapan Data Mining dengan MATLAB. Bandung: Rekayasa Sains
- YI Claudy, RS Pradana dan MA Fauzi. (2018). *Klasifikasi Dokumen Twitter Untuk Mengetahui Karakter Calon Karyawan Menggunakan Algoritma K-NN*. Teknologi Informasi dan Ilmu Komputer, Universitas Brawijaya
- Zarella, D. 2010. *The Social Media Marketing Book*. Jakarta: Serambi Ilmu Semesta.

LAMPIRAN

LAMPIRAN I : Tabel Daftar Slang-Words

No	Slang-Word	Kata Benar	No	Slang-Word	Kata Benar
1	/	atau	106	g	tidak
2	+	tambah	107	ga	tidak
3	=	sama dengan	108	gaada	tidak ada
4	abal	palsu	109	gabut	gaji buta
5	abal2	palsu	110	gada	tidak ada
6	abis	habis	111	gadapet	tidak dapat
7	abizz	habis	112	gapapa	tidak apa
8	acting	adegan	113	gadis	perempuan
9	action	aksi	114	gag	tidak
10	ad	ada	115	gaje	jelek
11	adek2	adik	116	gaje	aneh
12	aga	agak	117	gajebo	tidak jelas
13	agak	agak	118	gak	tidak
14	agr	agar	119	hadeuh	aduh
15	aj	saja	120	haha	senang
16	aja	saja	121	hater	pembenci
17	aje	saja	122	haters	pembenci
18	ak	aku	123	hati2	waspada
19	akn	akan	124	hbs	habis
20	akoh	aku	125	hbung	hubung
21	alay	norak	126	hbungan	hubungan
22	albm	album	127	hdap	hadap
23	alesan	alasan	128	ibu2	ibu-ibu
24	bgs	bagus	129	idup	hidup
25	bagus2	bagus	130	ig	instagram
26	bahas2	bahas	131	ikt	ikut
27	bpk	bapak	132	ikut2an	mengikuti
28	banget	sangat	133	ilang	hilang
29	bapk	bapak	134	ilangin	hilangkan
30	baper	bawa perasaan	135	ilke	like
31	bapuk	rusak	136	indo	indonesia
32	bwa	bawa	137	jdi	jadi
33	bbrp	beberapa	138	jdnya	jadinya
34	bcok	bacok	139	jdwal	jadwal
35	bcr	bicara	140	jealas	jelas
36	bda	beda	141	jg	juga
37	bdya	budaya	142	jgan	jangan
38	bebeb	pacar	143	jgn	jangan
39	begin	awal	144	jhat	jahat
40	bego	bodoh	145	jht	jahat
41	bejibun	banyak	146	jl	jalan
42	bener	benar	147	kagak	tidak
43	bnr	benar	148	kalo	kalua
44	ckp	cukup	149	kamseupay	kampung
45	cm	cuma	150	karn	karena
46	cmiiw	koreksi saya	151	karna	karena
47	cnderung	cenderung	152	kartun2	kartun
48	cnt	cinta	153	katain	mengatakan
49	congrats	selamat	154	kayak	seperti
50	cont	contoh	155	kbijakan	kebijakan
51	cr	cara	156	kbr	kabar
52	carany	caranya	157	lebeh	berlebihan
53	care	peduli	158	lebi	lebih
54	cartoon	kartun	159	legnda	legenda
55	catet	catet	160	lg	lagi
56	cbol	cebol	161	lgi	lagi
57	cepat2	cepat	162	lgkp	lengkap
58	cinta2an	cinta	163	lgkungan	lingkungan
59	ckap	cakap	164	lgnda	legenda

60	carany	caranya	165	lgs	langsung
61	cra	cara	166	lebeh	berlebihan
62	critan	ceritanya	167	lebi	lebih
63	critanya	ceritanya	168	legnda	legenda
64	d	di	169	lg	lagi
65	da	ada	170	lgi	lagi
66	dafuk	kurang ajar	171	ma	sama
67	dah	sudah	172	maaci	terimakasih
68	damat	tidak peduli	173	mabok	mabuk
69	dambil	diambil	174	maen	main
70	dapet	dapat	175	mager	malas gerak
71	dbiarkn	dibiarkan	176	maharani	mahal
72	ddik	didik	177	maho	homoseksual
73	dear	kepada	178	mai:	mayat
74	deh	kalua	179	mainstream	sangat umum
75	delete	hapus	180	maksh	terimakasih
76	dempa	demi apa	181	nice:bagus	bagus
77	dftar	daftar	182	nich	ini
78	dftr	daftar	183	nie	ini
79	dgang	dagang	184	objk	objek
80	dgn	dengan	185	obyek	objek
81	dhati	dihati	186	official	resmi
82	eko	ekonomi	187	olh	oleh
83	eksprsf	ekspresif	188	prbedaan	perbedaan
84	elo	kamu	189	prbedaan	perbedaan
85	elu	kamu	190	rmh	rumah
86	emak	ibu	191	rmja	remaja
87	emang	memang	192	sbgi	sabagai
88	emg	memang	193	sblm	sebelum
89	emng	memang	194	tetep	tetap
90	engga	tidak	195	tgl	tanggal
91	enjoy	Nyaman	196	udhlah	sudah
92	entar	Sebentar	197	umur	usia
93	eps	Episode	198	ungkp	ungkap
94	epsd	Episode	199	unyu	menggemaskan
95	faham	Paham	200	usahh	iya
96	fake	Falsu	201	utk	untuk
97	fans	penggemar	202	vacum	berhenti
98	faq	pertanyaan umum	203	vdeo	video
99	fham	Paham	204	wa	whatsapp
100	fhm	Paham	205	wanita	perempuan
101	fikir	Pokir	206	whyy	kenapa
102	filem	Film	207	x	kali
103	fis	Fisika	208	ykn	yakin
104	fix	Pasti	209	yng	yang
105	flm	Film	210	yuk	ayo

LAMPIRAN II : Tabel Daftar *StopWords-Removal*

No	<i>StopWord-Removal</i>	No	<i>StopWord -Removal</i>	No	<i>StopWord-Removal</i>	No	<i>StopWord-Removal</i>
1	-	31	dia	61	lain	91	sedangkan
2	:(32	dimana	62	maka	92	seharusnya
3	:)	33	dll	63	mari	93	sehingga
4	:d	34	dsb	64	masih	94	sekitar
5	d	35	dst	65	melainkan	95	selagi
6	:d	36	dulunya	66	mengapa	96	selain
7	a	37	guna	67	menurut	97	sementara
8	ada	38	hal	68	mereka	98	sementara
9	adalah	39	hanya	69	namun	99	seolah
10	agak	40	harus	70	nanti	100	seperti
11	agar	41	ia	71	nggak	101	seraya
12	akan	42	ingin	72	oh	102	serta
13	amat	43	ini	73	ok	103	sesuatu
14	bagaimanapun	44	itu	74	oleh	104	sesudah
15	bagi	45	itulah	75	pada	105	setelah
16	bahwa	46	jika	76	para	106	seterusnya
17	begitu	47	juga	77	pasti	107	setiap
18	belum	48	juga	78	pula	108	setidaknya
19	bisa	49	kah	79	pun	109	sudah
20	boleh	50	kami	80	saat	110	supaya
21	dahulu	51	karena	81	saja	111	tanpa
22	dalam	52	ke	82	sambil	112	tapi
23	dan	53	kecuali	83	sampai	113	telah
24	dapat	54	kemana	84	saya	114	tentang
25	dari	55	kembali	85	sebab	115	tentu
26	daripada	56	kenapa	86	sebagai	116	terhadap
27	demi	57	kepada	87	sebelum	117	tetapi
28	demikian	58	ketika	88	sebetulnya	118	toh
29	dengan	59	kita	89	secara	119	tolong
30	di	60	lagi	90	lain	120	untuk
121	walau	122	ya	123	yaitu	124	yakni
125	yang						