

PROGRAM STUDI TEKNIK INFORMATIKA

**PENERAPAN TEXT MINING DALAM MENGANALISIS
KEPRIBADIAN PENGGUNA MEDIA SOSIAL DENGAN K-
NEAREST NEIGHBOR (STUDI KASUS : FACEBOOK)**

M PRAMADANI RIYANIS PUTRA

151420126

Skripsi ini diajukan sebagai syarat memperoleh gelar Sarjana Komputer



FAKULTAS ILMU KOMPUTER

UNIVERSITAS BINA DARMA

2020



M PRAMADANI RIYANIS PUTRA

151420126

**PENERAPAN TEXT MINING DALAM MENGANALISIS
KEPRIBADIAN PENGGUNA MEDIA SOSIAL DENGAN K-
NEAREST NEIGHBOR (STUDI KASUS : FACEBOOK)**

Skripsi ini diajukan sebagai syarat memperoleh gelar Sarjana Komputer

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS BINA DARMA

2020

HALAMAN PENGESAHAN

**PENERAPAN TEXT MINING DALAM MENGANALISIS
KEPRIBADIAN PENGGUNA MEDIA DENGAN SOSIAL DENGAN K-
NEAREST NEIGHBOR (STUDI KASUS : FACEBOOK)**

M PRAMADANI RIYANIS PUTRA

15.142.0126

**Telah diterima sebagai salah satu syarat untuk memperoleh gelar Sarjana
Komputer pada Program Studi Teknik Informatika**

Palembang, Febuari 2020

Fakultas Ilmu Komputer

Universitas Bina Darma

Dekan,



Dosen Pembimbing,

A handwritten signature in black ink, belonging to Kiky Rizky Nova Wardani.


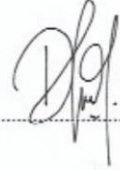

Kiky Rizky Nova Wardani, M.Kom

Dedy Syamsuar, S.Kom., M.I.T., Ph.D.

HALAMAN PERSETUJUAN

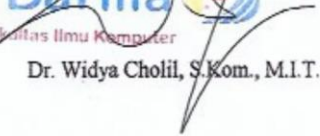
Skripsi Berjudul "PENERAPAN TEXT MINING DALAM MENGANALISIS KEPRIBADIAN PENGGUNA MEDIA SOSIAL DENGAN K-NEAREST NEIGHBOR (STUDI KASUS : FACEBOOK)" Oleh "M PRAMADANI RIYANIS PUTRA" telah dipertahankan di depan komisi penguji pada hari sabtu , tanggal 22 Februari 2020.

Komisi Penguji

1. Kiky Rizky Nova Wardani, M.Kom., Ketua (..... )
2. Diana, S.Si., M.Kom., Anggota (..... )
3. Nia Oktaviani, M.Kom., Anggota (..... )

Mengetahui,
Program Studi Teknik Informatika
Fakultas Ilmu Komputer
Universitas Bina Darma
Ketua Program Studi,

Universitas 
Bina Darma
Fakultas Ilmu Komputer


Dr. Widya Cholil, S.Kom., M.I.T.

SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : M Pramadani Riyanis Putra

NIM : 151420126

Dengan ini menyatakan bahwa:

1. Karya tulis saya (tugas skripsi) ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik baik sarjana di Universitas Bina Darma atau perguruan tinggi lain.
2. Karya tulis ini murni gagasan, rumusan dan penelitian saya sendiri dengan arahan tim pembimbing.
3. Didalam karya tulis ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dikutip dengan mencantumkan nama pengarang dan memasukkan ke dalam daftar rujukan.
4. Saya bersedia tugas skripsi, yang saya hasilkan diccek keasliannya menggunakan *turnitin* serta diunggah ke internet, sehingga dapat diakses publik secara daring.
5. Surat pernyataan ini saya tulis dengan sungguh – sungguh dan apabila terbukti melakukan penyimpangan atau ketidak benaran dalam pernyataan ini, maka saya bersedia menerima sanksi sesuai dengan peraturan dari perundang-undangan yang terbukti.

Demikian surat pernyataan ini saya buat agar dapat dipergunakan sebagaimana mestinya.

Palembang, Maret 2020

Yang membuat pernyataan




M Pramadani Riyanis Putra

151420126

MOTTO DAN PERSEMBAHAN

MOTTO :

- ❖ Positive thinking dalam segala hal niscaya alam semesta akan memberimu energi yang positive dalam melakukan hal yang positive dan akan berakhir dengan hasil yang positive.
- ❖ Sesungguhnya bersama kesulitan ada kemudahan, maka apabila engkau telah selesai (dari sesuatu urusan), tetaplah bekerja keras (untuk urusan yang lain). Dan hanya kepada Tuhanmulah engkau berharap (QS. Al-Insyirah, 6-8)
- ❖ Hiduplah layaknya hari ini, hari terakhir kamu hidup di dunia.

PERSEMBAHAN :

- ✓ Allah S.W.T. atas segala kenikmatan dan ridha-NYA serta memberikan kemudahan kepada saya dalam menyelesaikan skripsi ini.
- ✓ Kedua Orang Tua (Mama & Papa), terimakasih untuk semua pengorbanan dan jerih payah kalian dalam mendidik dan membesarkan saya hingga saya berada dititik ini.
- ✓ Kepada teman-teman Bromance yang selalu support (Bro Maik, Bro Joo, Bro Vins, Bro Tamagoci, Bro Woyo)

KATA PENGANTAR

Puji syukur kehadirat Allah SWT karena berkat rahmat dan karunia- Nya jualah, penelitian ini dapat diselesaikan guna memenuhi salah satu syarat untuk memperoleh gelar sarjana sebagai proses akhir dalam menyelesaikan pendidikan dibangku kuliah.

Dalam skripsi ini, tentunya masih jauh dari sempurna. Hal ini dikarenakan keterbatasnya pengetahuan yang dimiliki. Oleh karena itu dalam rangka melengkapi kesempurnaan dari penulisan skripsi ini diharapkan adanya saran dan kritik yang diberikan bersifat membangun.

Pada kesempatan yang baik ini, tak lupa penulis menghaturkan terima kasih kepada semua pihak yang telah memberikan bimbingan, pengarahan, nasehat dan pemikiran dalam penulisan skripsi ini, terutama kepada :

1. Dr. Sunda Ariana, M.Pd., M.M., selaku Rektor Universitas Bina Darma Palembang.
2. Dedi Syamsuar, S.Kom., M.IT., Ph.D., selaku Dekan Fakultas Ilmu Komputer
3. Dr. Widya Cholil, S.Kom., M.I.T., selaku Ketua Program Studi Informatika.
4. Kiki Rizky Nova Wardani., M.Kom., selaku Pembimbing yang telah memberikan bimbingan penulisan skripsi ini.
5. Orang Tua saya, ibu Ineke dan bapak Riyanis yang selalu memberikan dorongan dan masukan serta bantuan baik moril maupun materil yang tak ternilai harganya.
6. Saudaraku Novriandani serta keluarga besarku yang selalu mendukung pada setiap keadaan.
7. Keluarga besar Bromance angkatan 2019
8. Teman-teman seperjuangan Teknik Informatika 2015 Universitas Bina Darma

Palembang, Februari 2020

Penulis

ABSTRAK

Facebook merupakan aplikasi jejaring sosial dimana para pengguna mengungkapkan banyak tentang diri mereka sendiri melalui laman postingan mereka. Sehingga penulis ingin mengetahui informasi apa saja yang dapat diambil tentang kepribadian pengguna. Data mining memainkan peran penting yang bertujuan untuk mengubah data mentah menjadi suatu struktur yang dapat dimengerti untuk dapat digunakan lebih lanjut. *Text mining* mengacu pada proses mengambil informasi berkualitas tinggi dari teks, salah satu metode klasifikasi yang dapat digunakan ialah algoritma *K-Nearest Neighbor*. Berdasarkan teori *big five personality* hasil penelitian menyimpulkan bahwa tingkat akurasi yang diperoleh yaitu 92.92%, dari 550 data dengan nilai karakter kepribadian *openness* yang paling tinggi yaitu 239, *Conscientiousness* sebanyak 16 data, *Extraversion* sebanyak 173 data, *Agreeableness* sebanyak 50 data, *Neuroticism* sebanyak 33 data dan 39 data yang tidak dapat diklasifikasikan.

Kata Kunci : *Text Mining, Big Five Personality, K-Nearest Neighbor, Facebook*

ABSTRACT

Facebook is a social networking application where users reveal a lot about themselves through their posting pages. So the writer wants to know what information can be taken about the user's personality. Data mining plays an important role which aims to transform raw data into a structure that can be understood for further use. Text mining refers to the process of retrieving high quality information from text ,one of the classification methods that can be used is the K-Nearest Neighbor algorithm. Based on the theory of big five personality the results of the study concluded that the accuracy rate obtained was 92.92%, from 550 data with the highest openness personality character value of 239, Conscientiousness of 16 data, Extraversion of 173 data, Agreeableness of 50 data, Neuroticism of 33 data and 39 data that cannot be classified.

Keywords : *Text Mining, Big Five Personality, K-Nearest Neighbor, Facebook*

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
SURAT PERNYATAAN	iii
MOTTO DAN PERSEMBAHAN	iv
KATA PENGANTAR	v
ABSTRAK	vi
DAFTAR ISI	viii
DAFTAR GAMBAR	x
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan dan Manfaat Penelitian	4
1.4.1 Tujuan Penelitian	4
1.4.2 Manfaat Penelitian	4
1.5 Metodologi penelitian	5
1.5.1 Waktu Penelitian	5
1.5.2 Data Penelitian	5
1.5.3 Metode Penelitian	5
1.5.4 Metode Pengumpulan Data	5
1.5.4.1 <i>Web Scrapping</i>	5
1.5.5 Teknik Analisis	6
1.6 Sistematika Penulisan	7
BAB II TINJAUAN PUSTAKA	
2.1 Media Sosial Facebook	9
2.2 Text Mining	10
2.2.1 Tahapan <i>Text Mining</i>	12
2.2.2 Proses <i>Pro – Processing</i>	13
2.2.3 Proses <i>Cleaning Data</i>	14
2.3 K –Nearest Neighbor	14
2.4 Analisa Kepribadian	17
2.5 Penelitian Sebelumnya	19
2.5.1 Penelitian Agnes Theresa Damanik (2015)	20
2.5.2 Penelitian Tommy Tandra (2017)	20
2.5.3 Penelitian Kemas Muslim (2017)	21
2.6 Kerangka Berpikir	22

BAB III ANALISIS *TEXT MINING*

3.1 Proses awal <i>Text Mining</i>	24
3.1.1 Data.....	24
3.1.2 <i>Pre – processing</i>	25
3.1.3 <i>Data Cleaning</i>	27
3.1.4 <i>Feature Selection</i>	28
3.1.5 <i>Feature Extraction</i>	29

BAB IV HASIL DAN PEMBAHASAN

4.1 <i>Data Mining</i>	32
4.1.1 Implementasi dengan <i>RapidMiner</i>	32
4.2 Evaluasi	40

BAB V PENUTUP

1.1 Kesimpulan	41
1.2 Saran	42

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR GAMBAR

	Halaman
2.1 Gambar Tahap-Tahap <i>Text Mining</i>	12
2.2 Gambar Proses <i>Pre-processing</i>	13
2.3 Gambar Alur kerja <i>Cleaning Data</i>	14
2.4 Gambar Kerangka berfikir	22
3.1 Gambar Data Awal	24
3.2 Gambar Proses <i>Cleaning Data</i>	27
3.3 Gambar <i>Missing Values</i>	27
3.4 Gambar Hasil <i>Cleaning Data</i>	28
3.5 Gambar Data yang sudah diolah	29
3.6 Gambar Input Data Karakter	30
3.7 Gambar Status yang sudah di kategorikan	31
4.1 Gambar Input data ke <i>RapidMiner</i>	33
4.2 Gambar Hasil input data ke <i>Rapidminer</i>	33
4.3 Gambar <i>Replace missing values</i>	34
4.4 Gambar Hasil dari <i>Replace Missing Values</i>	34
4.5 Gambar <i>Select Attributes</i>	35
4.6 Gambar Hasil dari <i>Select Attributes</i>	36
4.7 Gambar <i>Cross Validation</i>	37
4.8 Gambar K-NN.....	37
4.9 Gambar <i>Apply Model</i>	38
4.10 Gambar Performace.....	39
4.11 Gambar Hasil Perhitungan	39

DAFTAR TABEL

	Halaman
2.1 Tabel kriteria-kriteria kepribadian	19
3.1 Hasil proses <i>Tokenizing</i>	25
3.2 Hasil Proses <i>Filtering</i>	26
3.3 Hasil Proses <i>Steaming</i>	26
3.4 Hasil Proses <i>Tagging</i>	26
3.5 Tabel kecocokan kepribadian.....	30