

**ANALISIS SENTIMEN PENGGUNA *TWITTER* TERHADAP  
PROGRAM DIGITALENT KOMINFO MENGGUNAKAN  
ALGORITMA *NAÏVE BAYES CLASSIFIER***



**SKRIPSI**

**FAISYAL FACHRUR ROSADI**

**1803040055**

**TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK DAN SAINS**

**UNIVERSITAS MUHAMMADIYAH PURWOKERTO**

**2023**

**ANALISIS SENTIMEN PENGGUNA *TWITTER* TERHADAP  
PROGRAM DIGITALENT KOMINFO MENGGUNAKAN  
ALGORITMA *NAÏVE BAYES CLASSIFIER***



**SKRIPSI**

diajukan untuk memenuhi salah satu syarat memperoleh gelar Sarjana  
Komputer

**FAISYAL FACHRUR ROSADI**

**1803040055**

**TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK DAN SAINS**

**UNIVERSITAS MUHAMMADIYAH PURWOKERTO**

**2023**

## HALAMAN PERSETUJUAN

Skripsi yang diajukan oleh:

Nama : Faisyal Fachrur Rosadi

NIM : 1803040055

Program Studi : Teknik Informatika

Fakultas : Teknik dan Sains

Perguruan Tinggi : Universitas Muhammadiyah Purwokerto

Judul : Analisis Sentimen Pengguna *Twitter* Terhadap Program

Digitalent Kominfo Menggunakan Algoritma *Naïve Bayes*

*Classifier*

Telah disetujui untuk diajukan dalam ujian skripsi

Purwokerto, 5 Januari 2023

PEMBIMBING

Maulida Ayu Fitriani, S.Kom., M.Cs.

NIK. 2160823


## HALAMAN PENGESAHAN

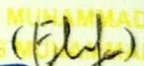
Skripsi yang diajukan oleh:

Nama : Faisyal Fachrur Rosadi  
NIM : 1803040055  
Program Studi : Teknik Informatika  
Fakultas : Teknik dan Sains  
Perguruan Tinggi : Universitas Muhammadiyah Purwokerto  
Judul : Analisis Sentimen Pengguna *Twitter* Terhadap Program Digitalent Kominfo Menggunakan Algoritma *Naïve Bayes Classifier*

telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Teknik Informatika, Fakultas Teknik dan Sains, Universitas Muhammadiyah Purwokerto.

### DEWAN PENGUJI

Penguji 1 (Pembimbing) : Maulida Ayu Fitriani, S.Kom., M.Cs. 

Penguji 2 : Elindra Ambar Pambudi, S.Kom., M.Kom. 

Penguji 3 : Harjono, S.T., M.Eng. 

Ditetapkan di : Purwokerto

Tanggal :

Mengetahui

Dekan Fakultas Teknik dan Sains

  
Ir. Teguh Marheni, S.T., ASEAN, Eng., ACPE., IPM

NIK. 2160172

## HALAMAN PERNYATAAN ORISINALITAS

Saya yang bertanda tangan dibawah ini:

Nama : Faisyal Fachrur Rosadi  
NIM : 1803040055  
Program Studi : Teknik Informatika  
Fakultas : Teknik dan Sains  
Perguruan Tinggi : Universitas Muhammadiyah Purwokerto

Menyatakan dengan sebenar-benarnya bahwa skripsi ini adalah hasil karya saya dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta bukan hasil penjiplakan dari karya orang lain.

Demikian pernyataan ini saya buat dan apabila kelak dikemudian hari terbukti ada unsur penjiplakan, saya bersedia mempertanggungjawabkan sesuai dengan ketentuan yang berlaku.

Purwokerto, 5 Januari 2023

Yang membuat pernyataan



Faisyal Fachrur Rosadi

## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika Universitas Muhammadiyah Purwokerto dan demi pengembangan ilmu pengetahuan, saya yang bertanda tangan di bawah ini:

Nama : Faisyal Fachrur Rosadi  
NIM : 1803040055  
Program Studi : Teknik Informatika  
Fakultas : Teknik dan Sains  
Perguruan Tinggi : Universitas Muhammadiyah Purwokerto  
Jenis Karya : Skripsi

Menyetujui untuk memberikan Hak Bebas Royalti Noneksklusif (*non-exclusive Royalty-Free Right*) kepada Universitas Muhammadiyah Purwokerto atas karya ilmiah saya yang berjudul:

### **Analisis Sentimen Pengguna *Twitter* Terhadap Program Digitalent Kominfo Menggunakan Algoritma *Naïve Bayes Classifier***

Beserta perangkat lunak yang ada (jika diperlukan), dengan Hak Bebas Royalti Noneksklusif ini Universitas Muhammadiyah Purwokerto berhak menyimpan, mengalih media/mengalih formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan skripsi saya dengan tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Purwokerto  
Pada tanggal : 5 Januari 2023  
Yang menyatakan,

Faisyal Fachrur Rosadi

## HALAMAN PERSEMBAHAN

Dengan segala kerendahan hati, serta rasa syukur terhadap Allah Subhanahu Wa Ta'ala yang telah memberikan rahmat dan nikmat-Nya, maka saya persembahkan Laporan Skripsi ini kepada:

1. Allah Subhanahu Wa Ta'ala yang senantiasa melimpahkan nikmat dan kasih sayang-Nya sehingga dapat menyelesaikan skripsi ini.
2. Kedua orang tua saya Bapak Sunarto dan Ibu Turyati yang senantiasa memberikan dukungan berupa moril, materil, bimbingan dan doanya untuk saya sampai saat ini.
3. Bapak Feri Wibowo, S.Kom., M.Cs. selaku Ketua Program Studi Teknik Universitas Muhammadiyah Purwokerto.
4. Ibu Maulida Ayu Fitriani, S.Kom., M.Cs. selaku Dosen Pembimbing Skripsi yang telah memberi arahan, bimbingan dan petunjuk dengan penuh kesabaran.
5. Terimakasih juga kepada teman Teknik Informatika dari tahun angkatan 2018 hingga saat ini yang telah memberi dukungan.
6. Sahabat-sahabat yang kusayangi Rifa Melati, Nadia Salsabila, David Saefullah, Nita Puspitasari, Okta Triana yang telah memberikan support, membantu memberikan ide dan perhatian dalam menemani pembuatan skripsi ini.

## MOTO

وَمَنْ يَتَّقِ اللَّهَ يَجْعَلْ لَهُ مِنْ أَمْرِهِ يُسْرًا

“Dan barangsiapa yang bertakwa kepada Allah, niscaya Allah menjadikan baginya kemudahan dalam urusannya”

[Ath-Thalaq/65 : 4]

*“The possibility of all those possibilities being possible is just another possibility that can possibly happen”*

**-Mark Lee-**

## KATA PENGANTAR

Puji syukur atas kehadiran Allah SWT yang telah melimpahkan rahmat, hidayah dan anugerah-Nya sehingga laporan skripsi ini dapat diselesaikan sesuai waktu yang telah ditentukan. Adapun judul skripsi ini adalah “**Analisis Sentimen Pengguna *Twitter* Terhadap Program Digitalent Kominfo Menggunakan Algoritma *Naïve Bayes Classifier***”. Skripsi ini diajukan untuk memenuhi syarat kelengkapan kurikulum pada program studi Teknik Informatika Fakultas Teknik dan Sains Universitas Muhammadiyah Purwokerto. Pembuatan laporan ini mengacu dan berpedoman pada bahan kuliah, petunjuk dari dosen pembimbing skripsi, referensi dan literatur yang terkait dengan penulisan laporan.

Laporan ini jauh dari kata sempurna dan masih banyak kekurangan mengingat keterbatasan pengalaman dan kemampuan, oleh karena itu kritik dan saran yang membangun sangat diharapkan demi hasil yang lebih baik dimasa mendatang. Akhirnya, besar harapan agar kehadiran laporan ini dapat memberikan manfaat yang berarti untuk para pembaca, dan yang terpenting adalah semoga dapat turut serta memajukan ilmu pengetahuan.

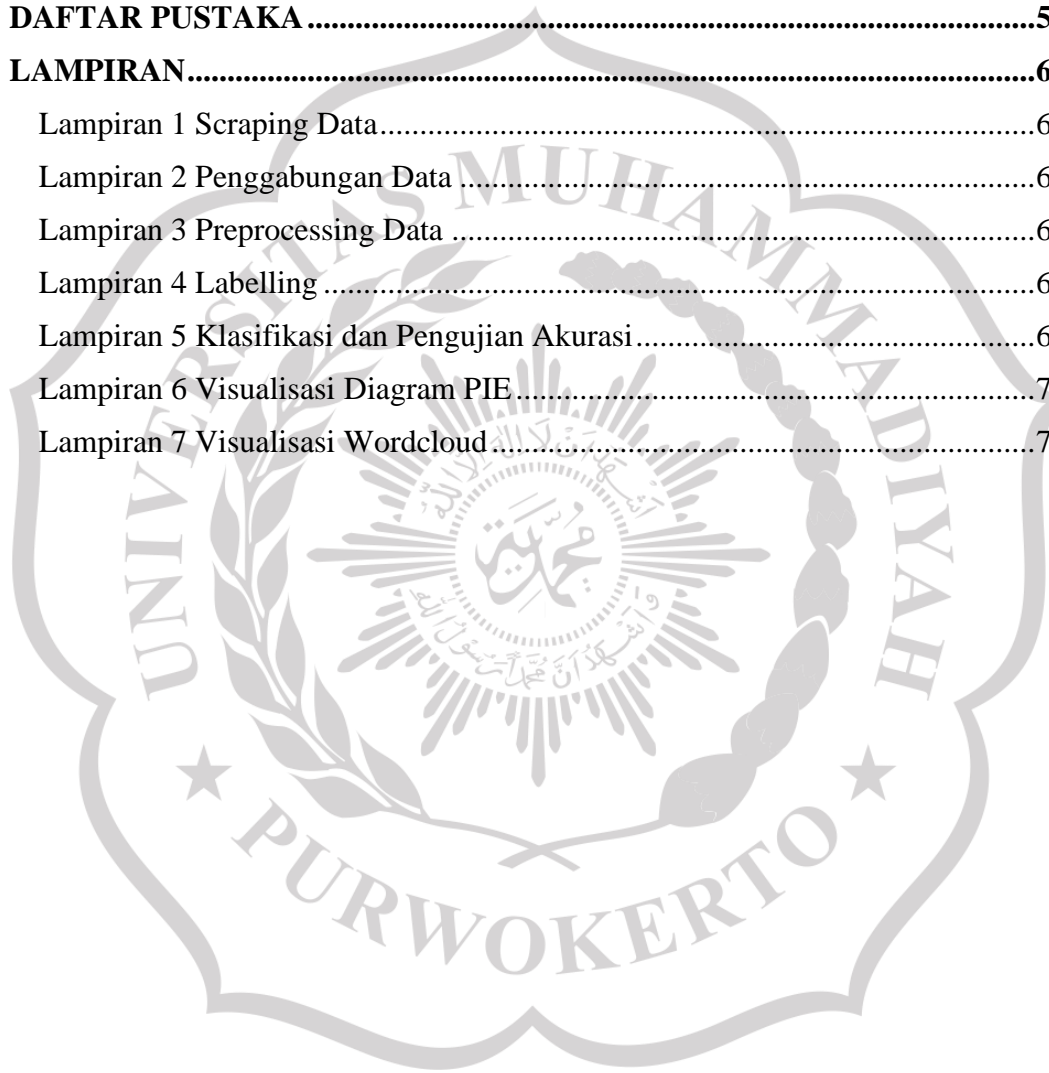
Penyusun,

Faisyal Fachrur Rosadi

## DAFTAR ISI

<b>HALAMAN SAMPUL</b> .....	<b>i</b>
<b>HALAMAN JUDUL</b> .....	<b>ii</b>
<b>HALAMAN PERSETUJUAN</b> .....	<b>iii</b>
<b>HALAMAN PENGESAHAN</b> .....	<b>iv</b>
<b>HALAMAN PERNYATAAN ORISINALITAS</b> .....	<b>v</b>
<b>HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI</b> .....	<b>vi</b>
<b>HALAMAN PERSEMBAHAN</b> .....	<b>vii</b>
<b>MOTO</b> .....	<b>viii</b>
<b>KATA PENGANTAR</b> .....	<b>ix</b>
<b>DAFTAR ISI</b> .....	<b>x</b>
<b>DAFTAR TABEL</b> .....	<b>xii</b>
<b>DAFTAR GAMBAR</b> .....	<b>xiii</b>
<b>INTISARI</b> .....	<b>xiv</b>
<b>ABSTRAK</b> .....	<b>xv</b>
<b>BAB I PENDAHULUAN</b> .....	<b>16</b>
A. Latar Belakang .....	16
B. Rumusan Masalah .....	19
C. Batasan Masalah.....	20
D. Tujuan Penelitian .....	20
E. Manfaat Penelitian .....	20
<b>BAB II TINJAUAN PUSTAKA</b> .....	<b>22</b>
A. Penelitian Terdahulu .....	22
B. Landasan Teori.....	28
<b>BAB III METODE PENELITIAN</b> .....	<b>34</b>
A. Jenis Penelitian.....	34
B. Metode Analisis Data.....	34
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	<b>43</b>
A. Pengumpulan Data .....	43
B. Langkah <i>Preprocessing</i> .....	44
C. Pelabelan Data.....	48

D. Implementasi Klasifikasi <i>Naïve Bayes</i> .....	48
E. Evaluasi <i>Confusion Matrix</i> .....	52
F. Visualisasi Data.....	54
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>58</b>
A. Kesimpulan .....	58
B. Saran.....	58
<b>DAFTAR PUSTAKA .....</b>	<b>59</b>
<b>LAMPIRAN.....</b>	<b>63</b>
Lampiran 1 Scraping Data.....	63
Lampiran 2 Penggabungan Data .....	63
Lampiran 3 Preprocessing Data .....	63
Lampiran 4 Labelling .....	66
Lampiran 5 Klasifikasi dan Pengujian Akurasi.....	68
Lampiran 6 Visualisasi Diagram PIE.....	72
Lampiran 7 Visualisasi Wordcloud.....	72

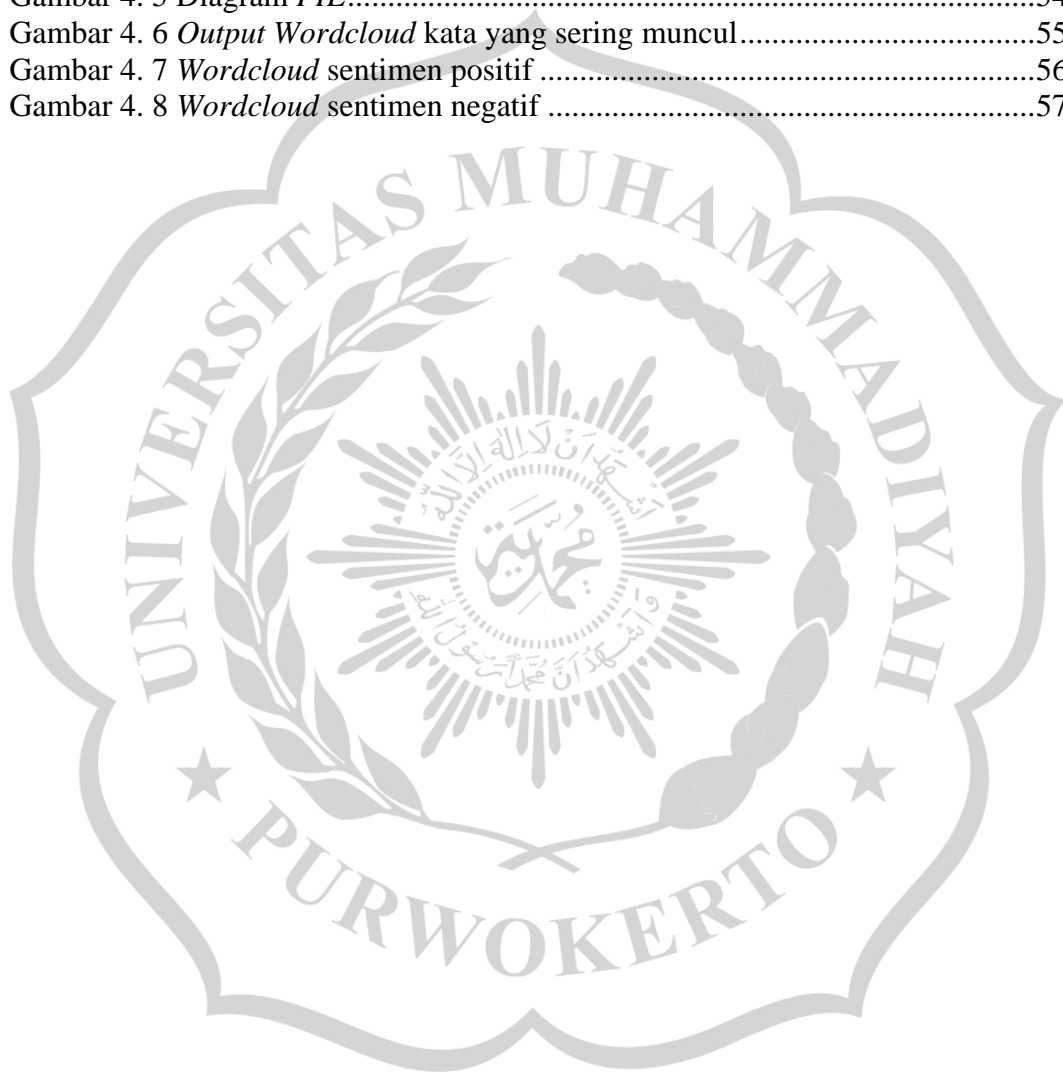


## DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu .....	22
Tabel 3. 1 <i>Cleaning Text</i> .....	35
Tabel 3. 2 <i>Tokenizing Text</i> .....	36
Tabel 3. 3 <i>Slangword</i> .....	36
Tabel 3. 4 <i>Stopword Removal</i> .....	37
Tabel 3. 5 <i>To Sentence</i> .....	37
Tabel 3. 6 <i>Stemming Text</i> .....	37
Tabel 4. 1 Proses <i>Scraping</i> .....	43
Tabel 4. 2 Proses <i>Cleaning</i> .....	45
Tabel 4. 3 Proses <i>Tokenizing</i> .....	45
Tabel 4. 4 Proses normalisasi kalimat .....	46
Tabel 4. 5 Proses <i>Filtering</i> .....	46
Tabel 4. 6 Proses <i>To Sentence</i> .....	47
Tabel 4. 7 Proses <i>Stemming</i> .....	47
Tabel 4. 8 Hasil Pelabelan Data .....	48
Tabel 4. 9 Hasil <i>CountVectorizer</i> .....	49
Tabel 4. 10 Hasil Data <i>Splitting</i> .....	49
Tabel 4. 11 pendekatan <i>Naïve Bayes Multinomial</i> .....	51
Tabel 4. 12 Sentimen Positif <i>Confusion Matrix</i> .....	53
Tabel 4. 13 Sentimen Negatif <i>Confusion Matrix</i> .....	53
Tabel 4. 14 Contoh <i>Tweet</i> Pada <i>Wordcloud</i> .....	55
Tabel 4. 15 Contoh <i>Wordcloud</i> Positif .....	56
Tabel 4. 16 Contoh <i>Wordcloud</i> Negatif .....	57

## DAFTAR GAMBAR

Gambar 3. 1 Metodologi Penelitian .....	34
Gambar 4. 1 <i>Splitting</i> data.....	50
Gambar 4. 2 <i>Output accuracy</i> .....	50
Gambar 4. 3 <i>Heatmap Confusion Matrix Naive Bayes</i> .....	52
Gambar 4. 4 <i>Output evaluasi classification report</i> .....	53
Gambar 4. 5 Diagram <i>PIE</i> .....	54
Gambar 4. 6 <i>Output Wordcloud</i> kata yang sering muncul.....	55
Gambar 4. 7 <i>Wordcloud</i> sentimen positif .....	56
Gambar 4. 8 <i>Wordcloud</i> sentimen negatif .....	57



## INTISARI

Sumber daya manusia yang berkualitas untuk menunjang kinerja produktifitas sangat penting, pemerintahan Indonesia melalui KOMINFO yang merupakan penanggung jawab bidang komunikasi dan informatika yang berfokus pada pembangunan sektor telekomunikasi, tatakelola internet dan digitalisasi memiliki program yang salah satu programnya untuk meningkatkan sumber daya manusia melalui program Digitalent. Program Digitalent merupakan program pelatihan pengembangan kompetensi sejak tahun 2018 dengan tujuan utamanya yaitu meningkatkan ketrampilan dan daya saing, produktifitas, profesionalisme sumber daya manusia bidang teknologi informasi dan komunikasi bagi angkatan kerja muda Indonesia, masyarakat umum dan aparatur negara, untuk itu perlu dilakukan penelitian guna mengetahui bagaimana sentimen publik terhadap program Digitalent apakah mayoritas publik menilai program Digitalent dengan sentimen positif atau menilai dengan sentimen negatif. Banyaknya jumlah data *review* membutuhkan waktu yang lama untuk membaca satu-persatu, maka dari itu tanggapan publik mengenai program Digitalent didapat dengan teknik *scraping* pada *Twitter* dari tanggal 1 Januari 2018 sampai tanggal 1 September 2022 dengan menghasilkan data sebanyak 9934. Penelitian ini *preprocessing* yang digunakan adalah *cleaning text*, *tokenizing text*, normalisasi, *filtering*, *to sentence*, *stemming text*. Klasifikasi *Naïve Bayes* yang digunakan yaitu klasifikasi teks *Naïve Bayes Multinomial* dan hasil ketepatan klasifikasi menggunakan *Naïve Bayes Classifier Multinomial* menghasilkan akurasi sebesar 79% dengan *precision* tertinggi 84%, *recall* tertinggi 89% dan *f1-score* tertinggi 86%.

Kata kunci: kominfo, digitalent, analisis sentimen, *naïve bayes classifier*.

## **ABSTRAK**

*Quality human resources to support productivity performance is very important, the Indonesian government through KOMINFO which is in charge of the communication and informatics sector which focuses on the development of the telecommunications sector, internet governance and digitalization has a program which one of its programs is to improve human resources through the Digitalent program. The Digitalent Program has been a competence development training program since 2018 with the main objective of increasing skills and competitiveness, productivity, professionalism of human resources in the field of information and communication technology for Indonesia's young workforce, the general public and the state apparatus, for this reason it is necessary to conduct research to find out how is the public sentiment towards the Digitalent program, does the majority of the public assess the Digitalent program with positive sentiments or judge it with negative sentiments. The large amount of review data requires a long time to read one by one, therefore public response regarding the Digitalent program was obtained by scraping techniques on Twitter from January 1 2018 to September 1 2022 by producing 9934 data. This research used preprocessing cleaning text, tokenizing text, normalization, filtering, to sentence, stemming text. The Naïve Bayes classification used is the Naïve Bayes Multinomial text classification and the results of the classification accuracy using the Naïve Bayes Classifier Multinomial produce an accuracy of 79% with the highest precision of 84%, the highest recall of 89% and the highest f1-score of 86%.*

*Keywords: kominfo, digitalent, sentiment analysis, naïve bayes classifier.*