

**KLASIFIKASI KEPERIBADIAN DISC BERDASARKAN  
CAPTION INSTAGRAM BAHASA INDONESIA MENGGUNAKAN  
ALGORITMA SUPPORT VECTOR MACHINE**

**SKRIPSI**



disusun oleh

**Bagas Wahyu Ikhsandi**

**18.12.0776**

**PROGRAM SARJANA  
PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2021**

**KLASIFIKASI KEPERIBADIAN DISC BERDASARKAN  
CAPTION INSTAGRAM BAHASA INDONESIA MENGGUNAKAN  
ALGORITMA SUPPORT VECTOR MACHINE**

**SKRIPSI**

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana  
pada Program Studi Sistem Informasi



disusun oleh

**Bagas Wahyu Ikhsandi**

**18.12.0776**

**PROGRAM SARJANA  
PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2021**

## **PERSETUJUAN**

### **SKRIPSI**

#### **KLASIFIKASI KEPERIBADIAN DISC BERDASARKAN CAPTION INSTAGRAM BAHASA INDONESIA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE**

Yang dipersiapkan dan disusun oleh :

**Bagas Wahyu Ikhsandi**

**18.12.0776**

Telah disetujui oleh Dosen Pembimbing Skripsi  
Pada tanggal 19 November 2021

**Dosen Pembimbing,**

**Anggit Dwi Hartanto, M.Kom**

**NIK. 190302163**

**PENGESAHAN**  
**SKRIPSI**  
**KLASIFIKASI KEPERIBADIAN DISC BERDASARKAN**  
**CAPTION INSTAGRAM BAHASA INDONESIA MENGGUNAKAN**  
**ALGORITMA SUPPORT VECTOR MACHINE**

Yang dipersiapkan dan disusun oleh

**Bagas Wahyu Ikhsandi**  
**18.12.0776**

Telah dipertahankan didepan Dewan Penguji  
Pada tanggal 19 November 2021

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Anggit Dwi Hartanto, M.Kom**

**NIK. 190302163**

**Hartatik, ST., M.Cs.**

**NIK. 190302232**

**Rizqi Sukma Kharisma, M.Kom**

**NIK. 190302215**

Skripsi ini telah diterima sebagai salah satu persyaratan  
Untuk memperoleh gelar Sarjana Komputer  
Tanggal 28 November 2021

**DEKAN FAKULTAS ILMU KOMPUTER**

**Hanif Al Fatta, S.Kom., M.Kom**

**NIK. 190302096**

## PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 19 November 2021



Bagas Wahyu Ikhsandi

NIM. 18.12.0776

## MOTTO

“ Sesungguhnya Allah tidak akan mengubah keadaan suatu kaum sebelum mereka  
mengubah keadaan diri mereka sendiri”

(QS. Ar- Ra'd [13]:11)

“One Good Think will lead to another”

-Fellexandro-

“Believe in yourself and all that you are. Know that there is something inside you  
that is greater than any obstacle”

-Christian D. Larson-

“Bermimpi aja dulu”

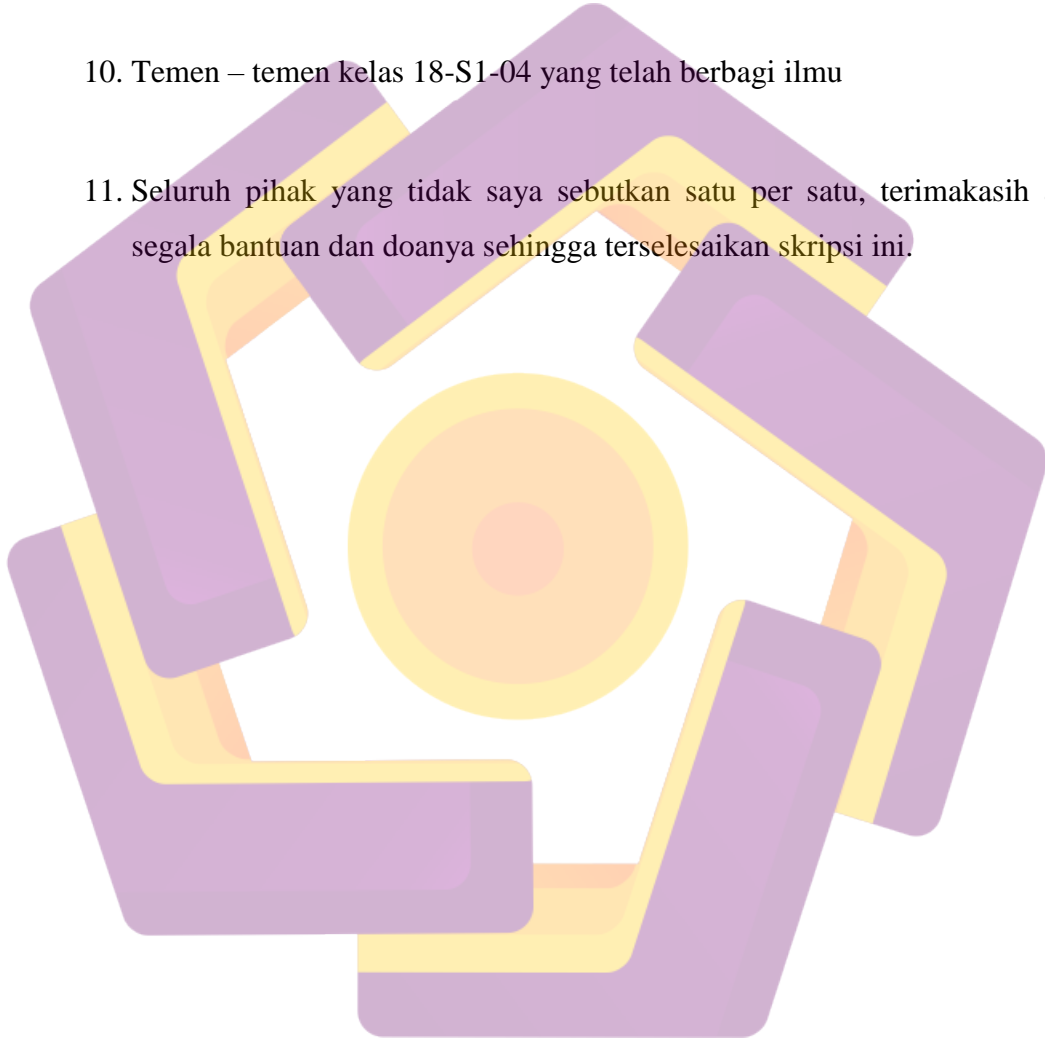
-Bagas Wahyu ikhsandi-

## PERSEMBAHAN

Alhamdulillah, segala puji bagi Allah SWT, yang telah memberikan rahmat, nikmat, dan hidayahnya sehingga saya bisa menyelesaikan skripsi ini sesuai dengan target dan mendapatkan hasil yang terbaik. Saya juga sangat berterima kasih kepada orang-orang yang telah secara langsung maupun tidak langsung yang telah membantu saya dalam menyelesaikan skripsi ini. Skripsi ini saya persembahkan kepada :

1. Kedua orang tua saya yang tak pernah lelah mendoakan, mendukung, membimbing, dan memberi semangat
2. Keluarga besar, yang selalu memberikan semangat dan nasehat
3. Beasiswa Unggulan yang mensupport saya secara finansial
4. Bapak Anggit Dwi Hartato, M.Kom. yang telah membimbing saya selama penelitian baik perlombaan dan skripsi.
5. Ahmad amdad Bariro, sebagai sahabat dan rekan seperjuangan saya selama perkuliahan dan sampai akhir penelitian yang selalu memberikan semua masukan dan berjuang sampai akhir.
6. Mas Bayu permana sejati sebagai mentor sekaligus role model, yang telah membimbing, memberi saran dan kritikan
7. Keluarga Besar Forum asisten teman - teman seperjuangan yang melakukan penelitian, Arbi, Ren, Yustika, Beta, Mitha, Anggita, Dwi, Lia, Dini, Iza, Handa, Nenden.

8. Teman – teman All in One, Febri, Ali, Wulan, Irma, yang selalu membantu, dan memberi masukan
9. Temen – temen kontrakan Tongkol, Rama, Bayu, Hasyim, dll, yang selalu membantu, dan berbagi canda tawa
10. Temen – temen kelas 18-S1-04 yang telah berbagi ilmu
11. Seluruh pihak yang tidak saya sebutkan satu per satu, terimakasih atas segala bantuan dan doanya sehingga terselesaikan skripsi ini.





## KATA PENGANTAR

Puji Syukur kita panjatkan kepada Allah atas berkat dan karunia-Nya sehingga penulis dapat menyelesaikan penulisan skripsi ini tepat pada waktunya dengan judul **“Klasifikasi Kepribadian Disc Berdasarkan Caption Instagram Bahasa Indonesia Menggunakan Algoritma Support Vector Machine”**.

Skripsi ini disusun untuk memenuhi salah satu persyaratan dalam menyelesaikan Program Strata-I Sistem Informasi di Universitas Amikom Yogyakarta. Selama mengikuti pendidikan Srtata-I Sistem Informasi sampai dengan proses penyelesaian skripsi, berbagai pihak telah memberikan fasilitas, membantu, membina, dan membimbing peulis untuk itu khususnya kepada:

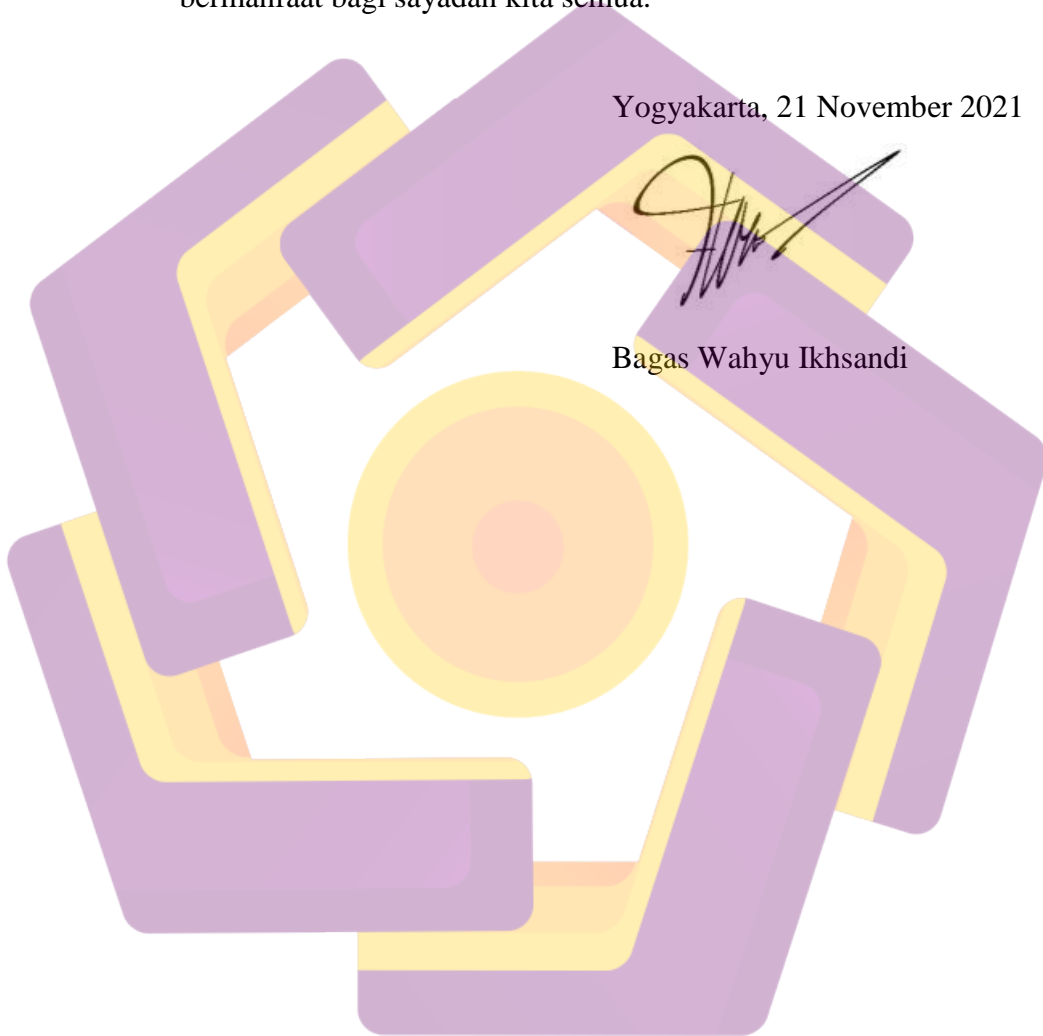
1. Bapak Prof. Dr. M. Suyanto , MM selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Bapak Anggit Dwi Hartanto, M.Kom. Selaku Kepala Program Studi Sistem Informasi Universitas Amikom Yogyakarta, sekaligus ddosen pembimbing saya yang selalu bijaksana memberikan bimbingan nasehat serta waktu nya selama penulisan skripsi ini.
4. Bapak dan Ibu Dosen Universitas Amikom Yogyakarta yang telah banyak memberikan ilmunya selama kuliah.
5. Keluarga Besar Pengurus Forum asisten Universitas Amikom Yogyakarta yang telah memberikan motivasi dan dorongan kepada saya.
6. Teman - teman 18 Sistem Informasi 4 yang sudah berjuang bersama – sama, membagi ilmu, canda tawa, dan kebersamaannya selama perkuliahan.
7. Semua pihak yang telah membantu baik secara langsung maupun tidak langsung.

Semoga Tuhan memberikan balasan yang lebih kepada semua yang telah ikut membantu saya hingga menyelesaikan skripsi ini. Demi perbaikan selanjutnya, saran dan kritik yang membangun diterima dengan senang hati dan rasa terima kasih. Semoga skripsi ini dapat bermanfaat bagi sayadan kita semua.

Yogyakarta, 21 November 2021



Bagas Wahyu Ikhsandi

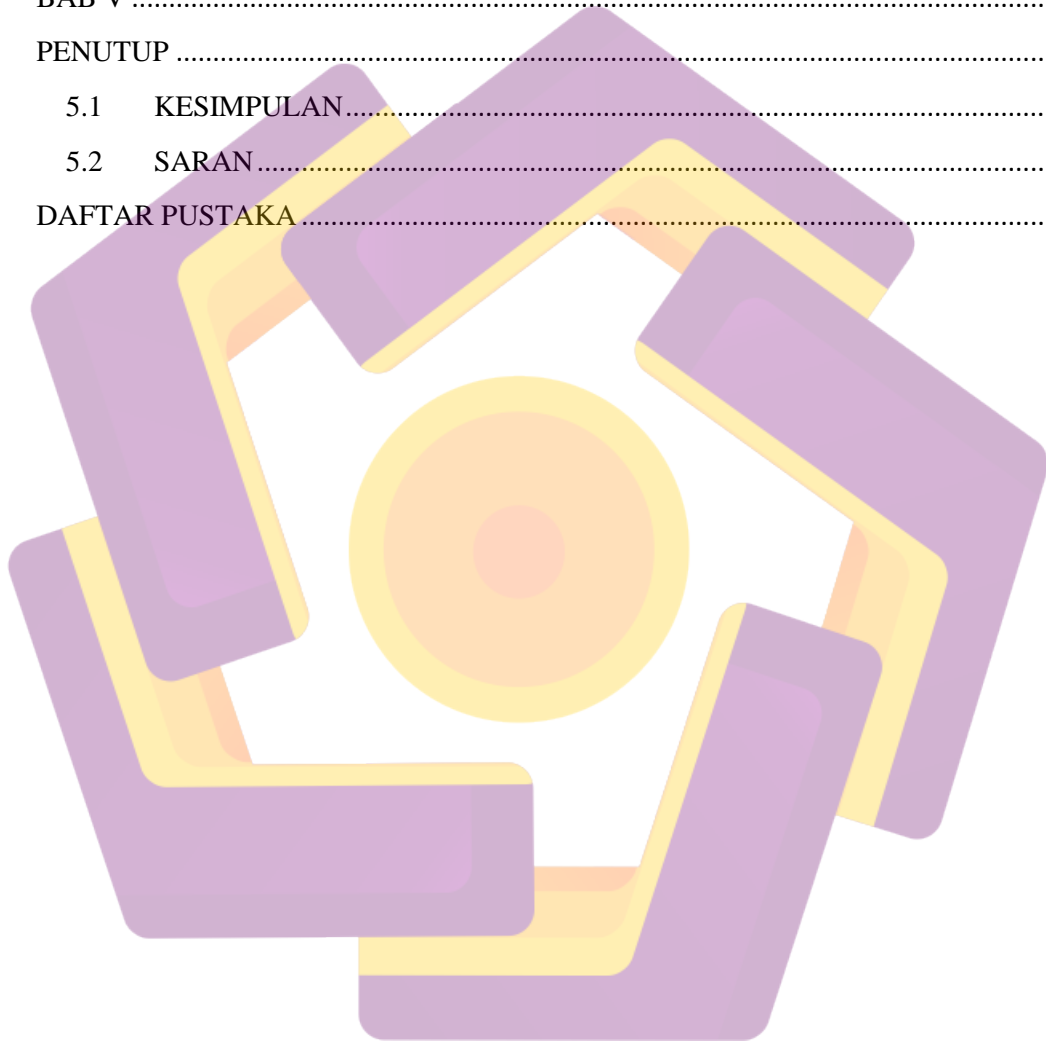


## DAFTAR ISI

SKRIPSI.....	i
PERSETUJUAN.....	ii
MOTTO.....	v
PERSEMBAHAN.....	vi
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xv
INTISARI.....	xvii
ABSTRACT.....	xviii
BAB I.....	1
PENDAHULUIAN.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	3
1.3. Batasan Masalah.....	3
1.4. Maksud dan Tujuan.....	4
1.5. Manfaat Penelitian.....	4
1.6. Metode Penelitian.....	5
1.6.1. Metode Studi Literatur.....	5
1.6.2. Metode Pengumpulan Data.....	5
1.6.3. Metode Analisis.....	5
1.6.4. Implementasi.....	5
1.6.5. Testing.....	5
1.7. Sistematika Penulisan.....	6
BAB II.....	7
LANDASAN TEORI.....	7
2.1. Tinjauan Pustaka.....	7
2.2. Landasan Teori.....	9

2.2.1. Text Mining .....	9
2.2.2. Natural Language Processing.....	9
2.2.3. Instagram.....	11
2.2.4 Support Vector Machine .....	11
2.2.5. Pembersihan Data .....	14
2.2.6. Term Frequency Inverse Document Frequency .....	15
2.2.7. Testing.....	16
2.2.8. Word Cloud.....	18
<b>BAB III .....</b>	<b>19</b>
<b>METODOLOGI PENELITIAN.....</b>	<b>19</b>
3.1. Alur metode Penelitian.....	19
3.2. Sumber Data dan Variabel Penelitian .....	19
3.1.1. Data Set.....	19
3.1.2. Variabel.....	20
3.1.3. Visualisasi Word Cloud .....	20
3.2. Validasi dan Pelabelan Data .....	24
3.4. Pre processing data .....	25
3.4.1. Cleaning data .....	25
3.4.2. Case Folding .....	26
3.4.3. Stopword removal .....	27
3.4.4. Tokenizing .....	28
3.5. Klasifikasi menggunakan support vector machine .....	28
3.5.1. Pembobotan kata dengan TF-IDF .....	28
3.5.2. Implementasi.....	32
3.5.3. Pengujian Confusion Matrix .....	36
3.5.4. Pengujian Performa.....	40
<b>BAB IV .....</b>	<b>41</b>
<b>HASIL DAN PEMBAHASAN.....</b>	<b>41</b>
4.1 Implementasi Sistem.....	41
4.1.1. Hardware dan software .....	41
4.1.2. Data Preprocessing.....	42

4.1.3.TF- IDF .....	47
4.1.4.Training.....	47
4.1.5.Testing.....	48
4.1.5.Hyperparameter Tuning .....	49
4.2. Perbandingan Antar Beberapa metode.....	50
BAB V .....	54
PENUTUP .....	54
5.1 KESIMPULAN.....	54
5.2 SARAN.....	54
DAFTAR PUSTAKA .....	55



## DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka.....	8
Tabel 2.2 Fungsi Karnael RBF .....	13
Tabel 2.3 confusion matrix .....	16
Tabel 3.1 Variabel Penelitian.....	20
Tabel 3.2 Sampel Data .....	25
Tabel 3.3 Cleaning Data.....	26
Tabel 3.4 Case Folding .....	26
Tabe 3.5 Stopword Removal.....	27
Tabel 3.6 Tokenizing .....	28
Tabel 3.7 Ilustrasi Sampel Data pada Perhitungan TF .....	29
Tabel 3.8 Ilustrasi Sampel Data pada Perhitungan DF .....	30
Tabel 3.9 Ilustrasi Sampel Data pada Perhitungan IDF.....	31
Tabel 3.10 Ilustrasi Sampel Data pada Perhitungan TF-IDF.....	31
Tabel 3.11 Ketepatan Klasifikasi SVM RBF $C=10, \gamma=1$ .....	33
Tabel 3.12 Fungsi Hyperplane SVM Kernel RBF .....	35
Tabel 3.13 confusion matrix .....	36
Tabel 3.14 False Positif.....	37
Tabel 3.15 Presisi.....	37

Tabel 3.16 False Negative..... 38

Tabel 3.17 Recall ..... 38

Tabel 4.1 Perbandingan antar beberapa metode ..... 52

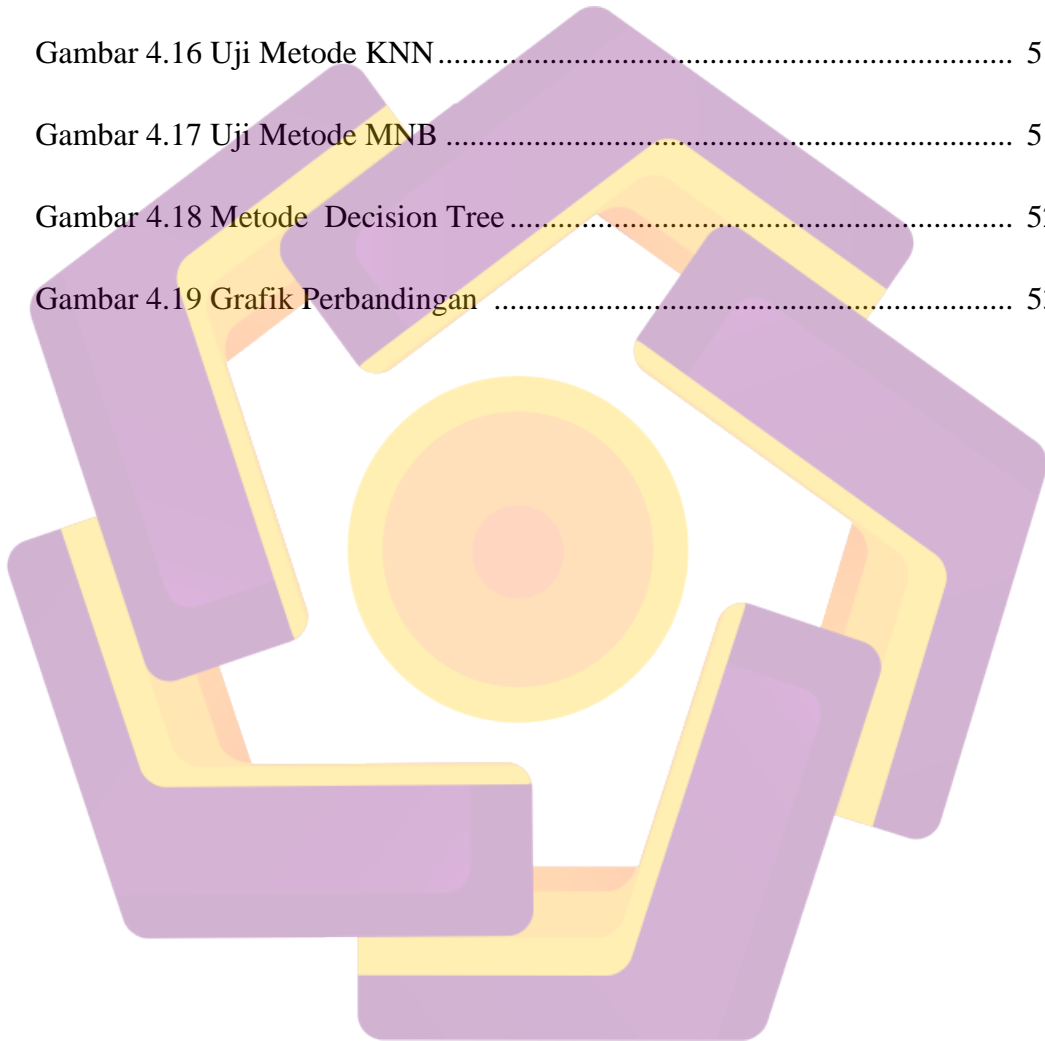


## DAFTAR GAMBAR

Gambar 2.1 Visualisasi Word Cloud .....	18
Gambar 3.1 Alur penelitian.....	19
Gambar 3.2. Sample Dataset.....	20
Gambar 3.3 Word Cloud Dominance.....	21
Gambar 3.4 Word Cloud Influence.....	22
Gambar 3.5 Word Cloud Steady.....	23
Gambar 3.6 Word Cloud Compliment.....	24
Gambar 3.7 Alur Preprocessing .....	25
Gambar 4.1 Proses import Data .....	42
Gambar 4.2 Pengambilan Data .....	43
Gambar 4.3 Download library.....	43
Gambar 4.4 Cleaning Data.....	44
Gambar 4.5 Case Folding.....	44
Gambar 4.6 Stopword Removal.....	45
Gambar 4.7 Tokenizing.....	46
Gambar 4.8 Preprocessing .....	46
Gambar 4.9 TF-IDF .....	47
Gambar 4.10 Training .....	48
Gambar 4.11 Testing.....	48



Gambar 4.12 Visualisasi confusion matrix .....	49
Gambar 4.13 Hyperparameter Tuning .....	49
Gambar 4.14 Parameter RBF .....	50
Gambar 4.15 Akurasi Hyperparameter Tuning.....	50
Gambar 4.16 Uji Metode KNN.....	51
Gambar 4.17 Uji Metode MNB .....	51
Gambar 4.18 Metode Decision Tree .....	52
Gambar 4.19 Grafik Perbandingan .....	53



## INTISARI

Perkembang media sosial instagram tiap tahunnya terus meningkat dan sangat diminati oleh masyarakat global terutama di Indonesia. Berdasarkan fungsinya aplikasi ini digunakan untuk berbagi foto maupun video pada sesama pengguna. Didalam instagram juga terdapat *caption* berupa *text* yang digunakan untuk mengungkapkan sebuah peristiwa, kejadian, ungkapan diri mereka sendiri dan hal yang sedang dirasakannya atau menggambarkan tentang diri mereka sendiri.

Seiring meningkatnya jumlah pengguna, maka meningkat pula jumlah postingan yang setiap harinya ayng tentu terdapat *caption*. Data *caption* tersebut akan digunakan sebagai sumber data penelitian untuk menganalisis profil kepribadian pengguna instagram berdasarkan *caption* bahasa indonesia dengan penilaian DISC (Dominance, influence, Steadiness, Compliance) dengan menggunakan algoritma Support Vector Machine.

Penelitian ini menghasilkan akurasi dari support vector machine sebesar 62% dan performa yang dihasilkan dari metode ini lebih baik dari metode Multinomial naïve bayes, KNN, dan Decision Tree.

**Kata kunci** : instagram, onfusion matrix, Support Vector Machine, DISC

## ABSTRACT

*The development of social media Instagram continues to increase every year and is in great demand by the global community, especially in Indonesia. Based on its function, this application is used to share photos and videos with other users. In Instagram there is also a caption in the form of text that is used to express an event, event, expression of themselves and what they are feeling or describe about themselves.*

*As the number of users increases, so does the number of posts that every day has a caption. The caption data will be used as a source of research data to analyze the personality profiles of Instagram users based on Indonesian captions with DISC (Dominance, Influence, Steadiness, Compliance) assessments using the Support Vector Machine algorithm.*

*This study resulted in an accuracy of 62% support vector machine and the resulting performance of this method is better than the Multinomial nave Bayes method, KNN, and Decision Tree.*

**Keywords:** *instagram, confusion matrix, Support Vector Machine, DISC*