

**IMPLEMENTASI ALGORITMA NAIVE BAYES TERHADAP KLASIFIKASI
SENTIMEN OPINI PENGGUNA INSTAGRAM PADA PRODUCT
ENDORSEMENT OLEH INFLUENCER DI INDONESIA**

SKRIPSI



disusun oleh

SALSABILA

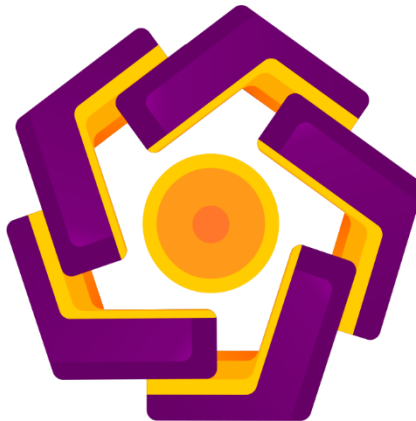
18.11.2107

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2022**

**IMPLEMENTASI ALGORITMA NAIVE BAYES TERHADAP KLASIFIKASI
SENTIMEN OPINI PENGGUNA INSTAGRAM PADA PRODUCT
ENDORSEMENT OLEH INFLUENCER DI INDONESIA**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

SALSABILA

18.11.2107

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2022**

PERSETUJUAN

SKRIPSI

**IMPLEMENTASI ALGORITMA NAIVE BAYES TERHADAP KLASIFIKASI
SENTIMEN OPINI PENGGUNA INSTAGRAM PADA PRODUCT
ENDORSEMENT OLEH INFLUENCER DI INDONESIA**

yang dipersiapkan dan disusun oleh

Salsabila
18.11.2107

Telah disetujui oleh Dosen Pembimbing Skripsi
Pada tanggal 22 Februari 2022

Dosen Pembimbing

Yoga Pristyanto S.Kom. M.Eng

NIK 190302412

PENGESAHAN

SKRIPSI

**IMPLEMENTASI ALGORITMA NAIVE BAYES TERHADAP KLASIFIKASI
SENTIMEN OPINI PENGGUNA INSTAGRAM PADA PRODUCT
ENDORSEMENT OLEH INFLUENCER DI INDONESIA**

yang dipersiapkan dan disusun oleh

Salsabila

18.11.2107

Telah dipertahankan didepan Dewan Penguji

Pada tanggal 22 Februari 2022

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Yoga Pristyanto, S.Kom. M.Kom

NIK. 190302412

Dina Maulina, M.Kom

NIK. 190302250

Kusnawi, S.Kom, M.Eng

NIK. 190302112

Skripsi ini telah diterima sebagai salah satu persyaratan

Untuk memperoleh gelar Sarjana Komputer

Tanggal 22 Februari 2022

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 diacy 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, 24 Maret 2022



Handwritten signature of Salsabila in black ink.

Salsabila

NIM. 18.11.2107

MOTTO

“Sesungguhnya Allah tidak akan mengubah keadaan suatu kaum sebelum mereka mengubah keadaan diri mereka sendiri. Dan apabila Allah menghendaki keburukan terhadap suatu kaum, maka tak ada yang dapat menolaknya dan tidak ada pelindung bagi mereka selain Dia”

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

“Jangan terlalu ambil hati dengan ucapan seseorang, kadang manusia punya mulut tapi belum tentu punya pikiran.”

-Albert Einstein-

“Jangan pergi mengikuti kemana jalan akan berujung. Buat jalanmu sendiri dan tinggalkanlah jejak.”

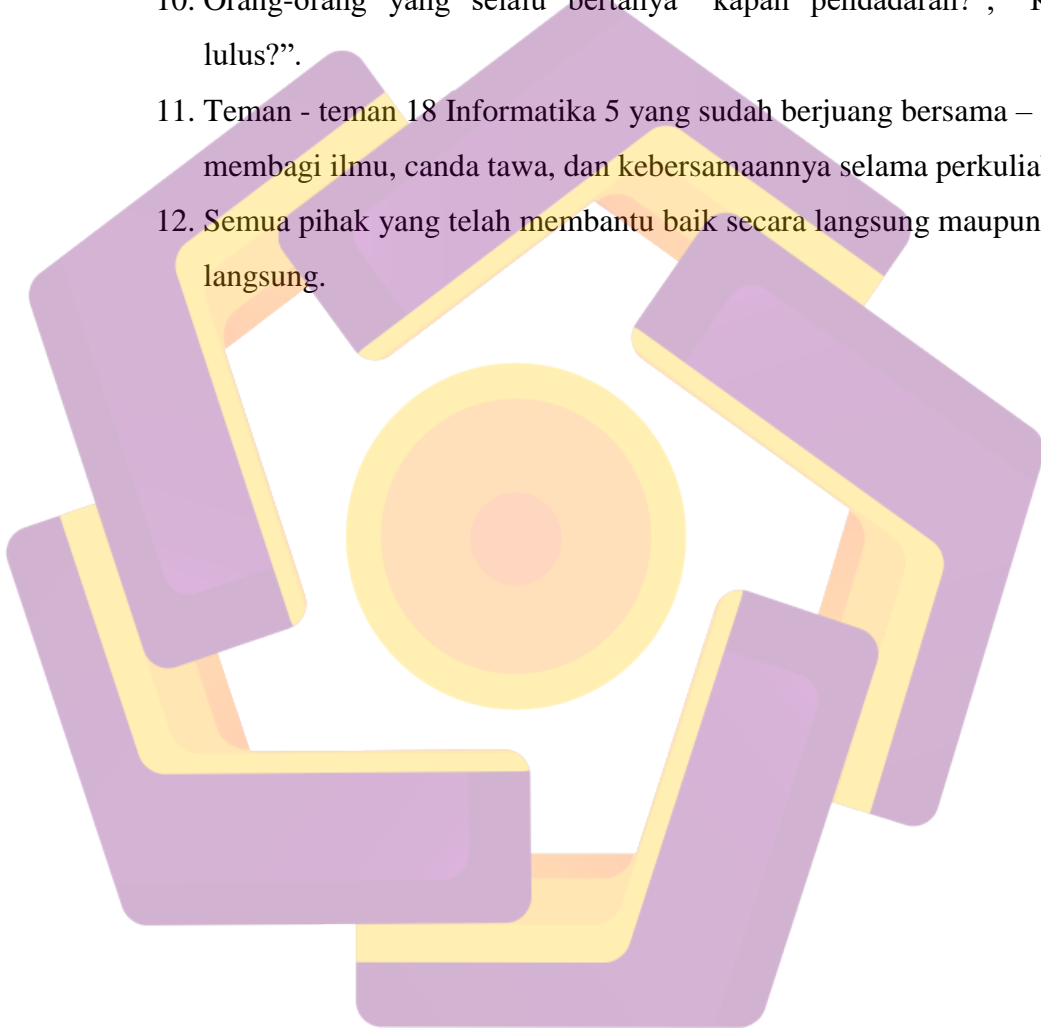
-Ralph Waldo Emerson-

PERSEMBAHAN

Segala puji dan syukur kehadirat Allah SWT atas berkat, rahmat dan hidayah-Nya sehingga penulis mampu menyelesaikan skripsi dengan baik. Penulis juga mengucapkan terima kasih kepada pihak – pihak yang telah berkontribusi baik secara langsung maupun tidak langsung baik dalam penelitian maupun dalam penyusunan naskah. Skripsi ini saya persembahkan kepada:

1. Allah SWT untuk segala nikmat yang luar biasa hambamu ini bersyukur atas nikmat yang telah diberikan.
2. Kepada dua orang hebat dalam hidup saya. Keduanya lah yang membuat segalanya menjadi mungkin sehingga saya bisa sampai pada tahap di mana skripsi ini akhirnya selesai. Terima kasih atas segala pengorbanan, nasihat dan doa baik yang tidak pernah berhenti kalian berikan kepadaku. Aku selamanya bersyukur dengan keberadaan kalian sebagai orangtua ku.
3. Last but not least, I wanna thank me, I wanna thank me for believing in me, I wanna thank me for doing all this hard work, I wanna thank me for having no days off, I wanna thank me for never quitting, I wanna thank me for just being me at all times.
4. Seluruh drama korea yang sudah menghibur saya selama ini.
5. DYAH AYU NUR WULAN, sebagai sahabat dan rekan seperjuangan saya selama di organisasi dan sampai akhir penelitian yang selalu memberikan semua masukan dan berjuang sampai akhir.
6. Teman – Teman Avengers, Wulan, Dicky, Isac, Luthfi, Delly yang telah membuat saya tumbuh dan selalu berbagi canda dan tawa bersama.
7. DEWI ARSHIETA LAKSMI dan QORINA LOIS FADILLA sebagai sahabat dan rekan seperjuangan saya selama perkuliahan dan mantan rekan satu kost yang telah banyak memberikan canda dan tawa.

8. Teman – Teman Kost sukun 24, Wulan, Dinda, Alin yang telah banyak membantu dan memberikan canda tawa.
9. Bagas Wahyu Ikhsandi, Ahmad amdad Bariro, junico sebagai mentor yang telah banyak membantu dalam pengerjaan skripsi ini.
10. Orang-orang yang selalu bertanya “kapan pendadaran?”, “Kapan lulus?”.
11. Teman - teman 18 Informatika 5 yang sudah berjuang bersama – sama, membagi ilmu, canda tawa, dan kebersamaannya selama perkuliahan.
12. Semua pihak yang telah membantu baik secara langsung maupun tidak langsung.



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 “Implementasi Algoritma Naive Bayes Terhadap Klasifikasi Sentimen Opini Pengguna Instagram Pada Product Endorsement Oleh Influencer Di Indonesia”.

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

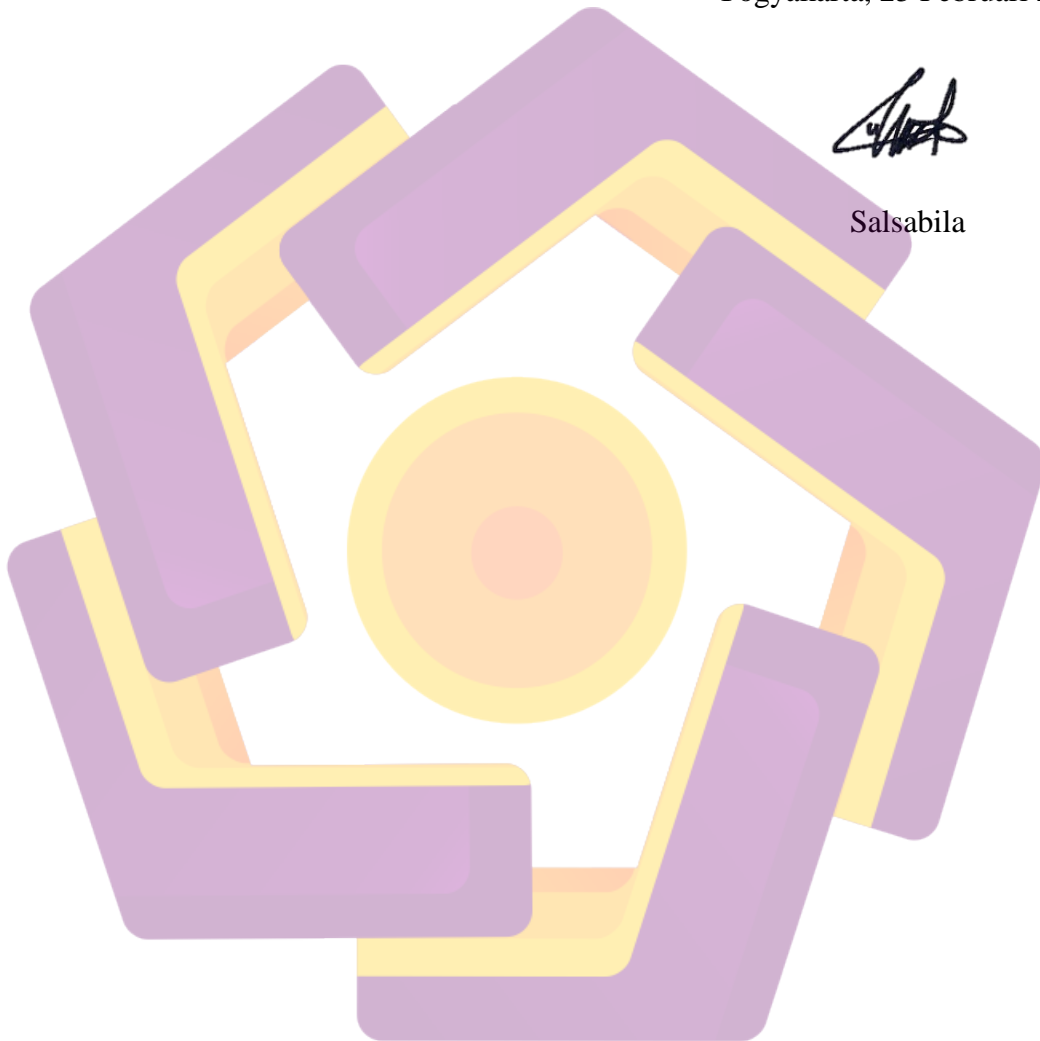
1. Bapak Prof. Dr. M. Suyanto , MM selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Bapak Yoga Pristyanto, S. Kom, M. Eng. Selaku dosen pembimbing saya yang telah dengan sabar dan tulus membimbing dan memberikan ilmu kepada penulis.
4. Bapak dan Ibu Dosen Universitas Amikom Yogyakarta yang telah banyak memberikan ilmunya selama kuliah.
5. Keluarga Besar Pengurus Amikom Computer Club (AMCC) Universitas Amikom Yogyakarta yang telah memberikan motivasi dan dorongan kepada saya.
6. Teman - teman 18 Informatika 5 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.

Penulis menyadari bahwa laporan skripsi ini masih jauh dari kata sempurna, untuk itu kritik dan saran yang membangun sangat diharapkan. Akhir kata semoga tulisan ini dapat memberikan manfaat bagi pembaca dan khususnya penulis sendiri.

Yogyakarta, 25 Februari 2022



Salsabila



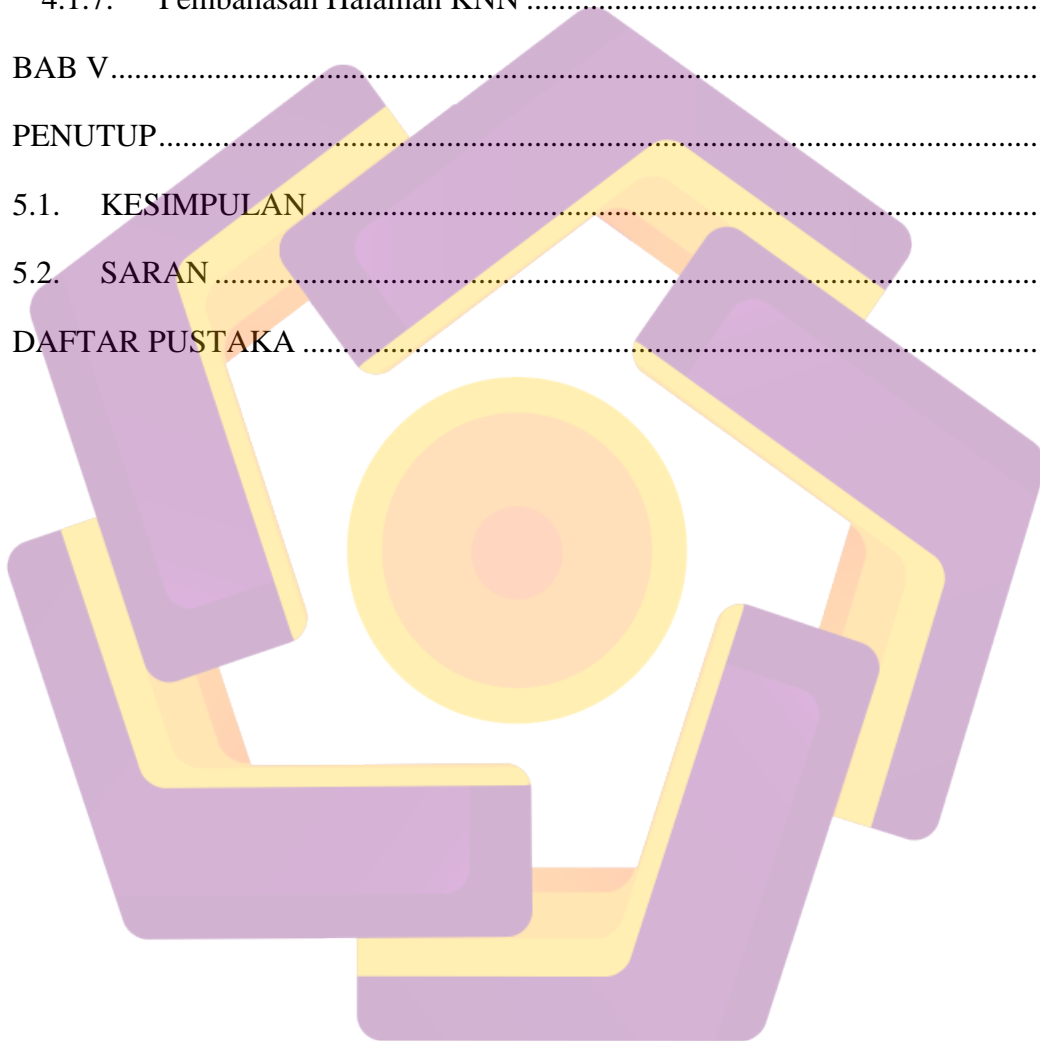
DAFTAR ISI

SKRIPSI.....	ii
MOTTO	iv
PERSEMBAHAN.....	vii
KATA PENGANTAR	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xv
DAFTAR TABEL.....	xvi
INTISARI.....	xviii
ABSTRAK.....	xix
BAB I.....	1
PENDAHULUAN	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah.....	4
1.4. Tujuan Penelitian	5
1.5. Manfaat Penelitian	5
1.6. Metode Penelitian.....	6
1.6.1. Metode Studi Literatur.....	6
1.6.2. Metode Pengumpulan Data.....	6
1.6.3. Metode Klasifikasi.....	6
1.6.4. Metode Testing	7

1.7. Sistematika Penulisan	7
BAB II.....	8
TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	8
2.1. Tinjauan Pustaka	8
2.2. Landasan Teori.....	14
2.2.1. Text Mining	14
2.2.2. Analisis Sentimen	15
2.2.3. <i>Instagram</i>	16
2.2.3.1. <i>Instagram Scraper</i>	16
2.2.4. <i>Influencer Marketing</i>	17
2.2.5. <i>Naive Bayes Classifier</i>	17
2.2.6. <i>Multinomial Naive Bayes Classifier</i>	20
2.2.7. <i>Text Preprocessing</i>	22
2.2.7.1. <i>Cleaning Data</i>	22
2.2.7.2. <i>Case Folding</i>	22
2.2.7.3. <i>Stopword removal</i>	22
2.2.7.4. <i>Tokenizing</i>	22
2.2.8. Pembobotan Kata (Term Weighting).....	23
2.2.9. <i>Testing</i>	24
2.2.9.1. <i>Confusion Matrix</i>	24
BAB III	24
METODOLOGI PENELITIAN.....	24
3.1. Alur metode Penelitian.....	24

3.2.	Sumber Data dan Variabel Penelitian	25
3.2.1.	Data Set.....	25
3.2.2.	Variabel.....	25
3.3.	Validasi dan Pelabelan Data	26
3.4.	<i>Pre processing data</i>	27
3.4.1.	<i>Cleaning data</i>	27
3.4.2.	<i>Case Folding</i>	28
3.4.3.	<i>Stopword Removal</i>	28
3.4.4.	<i>Tokenizing</i>	29
3.5.	Analisis Kebutuhan	30
3.5.1.	Hardware dan software	30
3.6.	Klasifikasi menggunakan Naive Bayes Classifier	31
3.6.1.	Pembobotan kata dengan TF-IDF.....	31
3.6.2.	Implementasi.....	34
3.5.2.1.	Klasifikasi Menggunakan <i>Multinomial Naive Bayes Classifier</i>	34
3.6.3.	Pengujian <i>Confusion Matrix</i>	41
3.6.4.	Pengujian Performa.....	45
BAB IV	43
HASIL DAN PEMBAHASAN	43
4.1.	Implementasi dan Pembahasan Sistem	43
4.1.1.	Pembahasan Halaman Dashboard.....	43
4.1.2.	Pembahasan Halaman Dataset	44
4.1.3.	Pembahasan Halaman Preprocessing.....	45

4.1.4.	Pembahasan Halaman Klasifikasi.....	46
4.1.5.	Pembahasan Halaman Pengujian MNB.....	47
4.1.6.	Pembahasan Halaman Decision Tree	48
4.1.7.	Pembahasan Halaman KNN	49
BAB V.....		52
PENUTUP.....		52
5.1.	KESIMPULAN.....	52
5.2.	SARAN	52
DAFTAR PUSTAKA		54



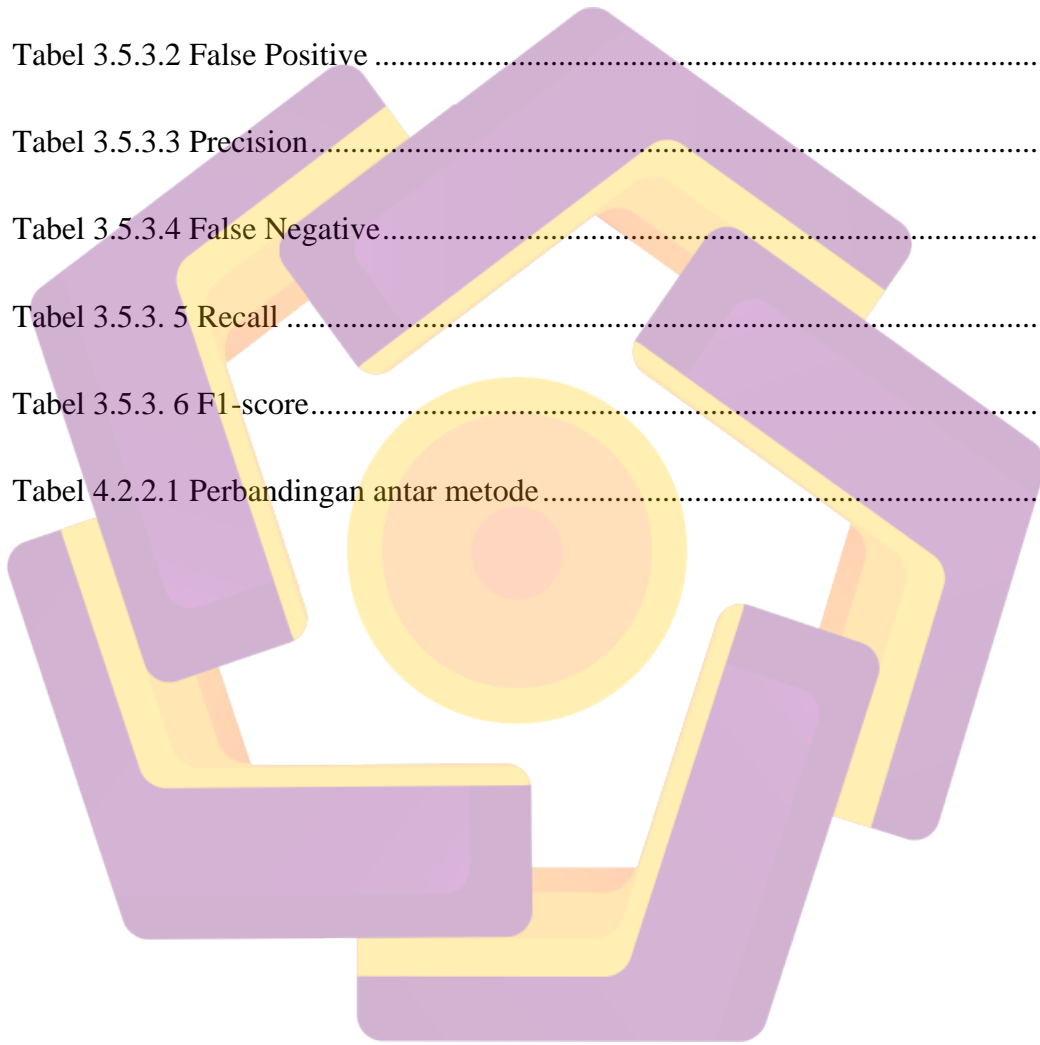
DAFTAR GAMBAR

Gambar 3. 1 Alur penelitian.....	24
Gambar 3.2.1 Sample Dataset.....	25
Gambar 3.4.1 Alur Preprocessing Data.....	27
Gambar 4.2.2.1 Halaman Dashboard.....	43
Gambar 4.2.2. 2 Halaman Dataset.....	44
Gambar 4.2.2.3 Halaman Text Processing.....	45
Gambar 4.2.2.4 Halaman Klasifikasi.....	46
Gambar 4.2.2.5 Halaman Pengujian MNB.....	47
Gambar 4.2.2.6 Halaman Decision Tree.....	48
Gambar 4.2.2.7 Halaman Pengujian KNN.....	49

DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka.....	11
Tabel 2.2.9.1 Confusion Matrix	24
Tabel 3.2.2.1 Variabel Penelitian.....	26
Tabel 3.3.1 Sample Data	26
Tabel 3.4.1.1 Proses Cleaning Data	27
Tabel 3.4.2.1 Proses Case Folding Data	28
Tabel 3.4.3.1 Proses Stopword Removal Data.....	29
Tabel 3.4.4.1 Proses Tekenizing Data.....	29
Tabel 3.5.1.1 Ilustrasi Sample Data Perhitungan TF	31
Tabel 3.5.1.2 Ilustrasi Sample Data Perhitungan DF.....	32
Tabel 3.5.1.3 Ilustrasi Sample Data Perhitungan IDF	32
Tabel 3.5.1.4 Ilustrasi Sample Data Perhitungan TF-IDF	33
Tabel 3.5.2.1.1 Contoh Kasus Data Training.....	34
Tabel 3.5.2.1.2 Hasil Perhitungan Prior Probabilities setiap kelas	35
Tabel 3.5.2.1.3 Contoh term pada dokumen	36
Tabel 3.5.2.1.4 Model Perhitungan Klasifikasi.....	37
Tabel 3.5.2.1.5 Data Testing	38

Tabel 3.5.2.1.6 Hasil Nilai Conditional Probabilitas	39
Tabel 3.5.2.1.7 Hasil Perkalian Nilai Conditional Probabilitas	40
Tabel 3.5.3.1 Visualisasi confusion matrix.....	41
Tabel 3.5.3.2 False Positive	42
Tabel 3.5.3.3 Precision.....	42
Tabel 3.5.3.4 False Negative.....	43
Tabel 3.5.3. 5 Recall	43
Tabel 3.5.3. 6 F1-score.....	44
Tabel 4.2.2.1 Perbandingan antar metode.....	50



INTISARI

Media sosial instagram merupakan salah satu media jejaring sosial yang dapat dimanfaatkan sebagai media pemasaran langsung. Media sosial instagram sebagai platform yang biasa digunakan untuk Strategi *Influencer Marketing* yang dimana akan sangat berpengaruh dan sangat penting bagi kelangsungan bisnis yang bergerak di dunia digital. Selain itu dengan penggunaan media sosial instagram ini masyarakat juga dengan mudah menyampaikan respon atau opini terhadap *Product Endorsement* yang dilakukan oleh pihak *influencer* tersebut. Berdasarkan fungsinya aplikasi ini digunakan untuk berbagi foto maupun video pada sesama pengguna. Didalam instagram juga terdapat komentar berupa text yang digunakan untuk merespon *Product Endorsement* yang telah diposting oleh pihak *influencer*.

Dengan melakukan analisis sentimen, pelaku bisnis dapat melihat seberapa besar pengaruh yang dihasilkan dari penggunaan jasa *endorsement* oleh *influencer*. Respon publik dari jasa yang dilakukan *influencer* terhadap produk atau brand apakah cenderung positif atau negatif.

Penelitian ini menggunakan dataset yang diambil langsung dari komentar instagram yang diambil dengan teknik scraping dan dijadikan sebagai objek penelitian. Data komentar dari Instagram akan menjadi alat pengukuran dan pengujian untuk membangun model sentiment analysis dengan bantuan metode Naive Bayes Classifier menghasilkan akurasi sebesar 94.62% dan performa yang dihasilkan dari metode ini lebih baik dari metode Decision Tree tapi tidak lebih baik dari metode KNN.

Kata Kunci : *Influencer Marketing, Instagram, Product Endorsement, Analisis sentimen, Naive Bayes Classifier.*

ABSTRAK

Instagram social media is one of the social networking media that can be used as a direct marketing medium. Instagram social media as a platform commonly used for Influencer Marketing Strategy which will be very influential and very important for the continuity of businesses that are engaged in the digital world. In addition, with the use of Instagram social media, the public also easily conveys responses or opinions on Product Endorsements carried out by the influencer. Based on its function, this application is used to share photos and videos with other users. In Instagram there are also comments in the form of text that are used to respond to Product Endorsements that have been posted by influencers.

By conducting sentiment analysis, business people can see how much influence is generated from the use of endorsement services by influencers. The public response from the services performed by influencers on the product or brand tends to be positive or negative.

This study uses a dataset taken directly from Instagram comments taken with the scraping technique and used as the object of research. Comments data from Instagram will be a measurement and testing tool to build a sentiment analysis model with the help of the Naive Bayes Classifier method resulting in an accuracy of 94.62% and the performance resulting from this method is better than the Decision Tree method but not better than the KNN method.

Keywords: *Influencer Marketing, Instagram, Product Endorsement, Sentiment Analysis, Naive Bayes Classifier.*