

**SENTIMEN ANALISIS PROVIDER BY.U
BERDASARKAN OPINI DI MEDIA SOSIAL TWITTER**

SKRIPSI



UNIVERSITAS AMIKOM YOGYAKARTA

Disusun oleh:

Thio Tangkas Kalyjaga Erz

17.11.1360

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2021

**SENTIMEN ANALISIS PROVIDER BY.U
BERDASARKAN OPINI DI MEDIA SOSIAL TWITTER**

SKRIPSI

Untuk memenuhi Sebagian persyaratan

mencapai derajat Sarjana S1

pada Program Studi Informatika



UNIVERSITAS AMIKOM YOGYAKARTA

Disusun oleh:

Thio Tangkas Kalyjaga Erz

17.11.1360

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2021

PERSETUJUAN

SKRIPSI

**SENTIMEN ANALISIS PROVIDER BY.U BERDASARKAN
OPINI DI MEDIA SOSIAL TWITTER**

yang dipersiapkan dan disusun oleh

Thio Tangkas Kalyjaga Erz

17.11.1360

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 10 April 2021

Dosen Pembimbing,

Sumarni Adi, S.Kom, M.Cs

NIK. 190302256

PENGESAHAN
SKRIPSI
SENTIMEN ANALISIS PROVIDER BY.U BERDASARKAN
OPINI DI MEDIA SOSIAL TWITTER

yang dipersiapkan dan disusun oleh

Thio Tangkas Kalyjaga Erz

17.11.1360

telah dipertahankan di depan Dewan Penguji

pada tanggal 21 April 2021

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Ika Nur Fajri, M.Kom

NIK. 190302268

Erni Seniwati, S.Kom, M.Cs

NIK. 190302231

Sumarni Adi, S.Kom, M.Cs

NIK. 190302256

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer

Tanggal 21 April 2021

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta, M.Kom

NIK. 190302096

PERNYATAAN

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 Mei 2021



Thio Tangkas Kalyaga Erz
NIM 17.11.1360

MOTTO

“Hidup adalah keberanian menghadapi tanda tanya.”



PERSEMBAHAN

Puji syukur saya panjatkan kehadiran Allah yang telah memberikan berkat yang luar biasa kepada saya, sehingga saya bisa menyelesaikan skripsi ini dengan baik. Saya juga sangat berterima kasih kepada orang-orang yang telah membantu saya secara langsung maupun tidak langsung dalam menyelesaikan skripsi ini. Skripsi ini saya persembahkan kepada :

1. Kedua orang tercinta Bapak Sisyanto dan Ibu Eli Rosmeri yang tidak pernah lelah mendoakan dan selalu memberikan support baik finansial maupun dalam dukungan lainnya. Terimakasih telah menjadi orang tua yang sangat istimewa.
2. Ibu Sumarni Adi, S.Kom, M.Cs selaku dosen pembimbing yang telah membimbing peneliti didalam pembuatan skripsi
3. Serta semua pihak yang telah membantu serta mendukung saya yang tidak bisa saya sebutkan satu persatu.

KATA PENGANTAR

Assalamua'laikum Warohmatullahi Wabarokatuhu

Alhamdulillah, puji syukur penulis panjatkan atas kehadiran Allah Yang Maha Esa yang telah memberikan segala rahmat dan hidayah-Nya sehingga penelitian dengan judul “SENTIMEN ANALISIS PROVIDER BY.U BERDASARKAN OPINI DI MEDIA SOSIAL TWITTER”. Dalam Penyusunannya, skripsi yang dibuat merupakan salah satu dari syarat kelulusan bagi setiap mahasiswa Universitas AMIKOM Yogyakarta dan menjadi sebuah bukti bahwa mahasiswa telah mampu menyelesaikan perkuliahan pada jenjang program Strata 1 dan untuk mendapat gelar Sarjana Komputer. Pada kesempatan kali ini penulis ingin mengucapkan terimakasih kepada:

1. Bapak Prof. Dr. M. Suyanto, MM selaku Rektor Universitas AMIKOM Yogyakarta .
2. Ibu Sumarni Adi, S.Kom, M.Cs selaku dosen pembimbing yang telah membimbing peneliti didalam pembuatan skripsi
3. Bapak dan Ibu Dosen Universitas AMIKOM Yogyakarta yang telah memberikan ilmu pengetahuan bagi peneliti dalam proses perkuliahan
4. Kedua Orang tercinta yaitu ibu Eli Rosmeri dan bapak Sisyanto serta kakak dan adik tersayang yaitu Tegar Al Aqbar Elrozis dan Tri Ageng Pangestu yang selalu mendoakan, memberikan semangat, perhatian dan bantuan dalam meyelesaikan skripsi ini.

5. Teman-teman baracas, squad sadyo, sobat jebara, dan sahabat-sahabtku semuanya yang tidak dapat saya sebutkan satu persatu yang telah memberikan dukungan dan motivasi agar penulis segera menyelesaikan skripsi ini.

6. Orang terkasih yang selalu mendukung kelancaran dalam pengerjaan skripsi ini.

Peneliti sendiri memiliki kesadaran bahwa pembuatan skripsi yang dilakukan masih sangat banyak kekurangan dan kelemahan. Oleh karena itu, peneliti berharap kepada semua pihak untuk dapat memberikan kritik dan masukan yang membangun guna menambah kesempurnaan skripsi ini. Dan peneliti tetap memiliki harapan bahwa skripsi yang dibuat memiliki manfaat bagi seluruh pihak yang membacanya. Wassalamu'alaikum Warahmatullahi Wabarakatuhu

Yogyakarta,..... 2021

Thio Tangkas Kalyjaga Erz
17111360

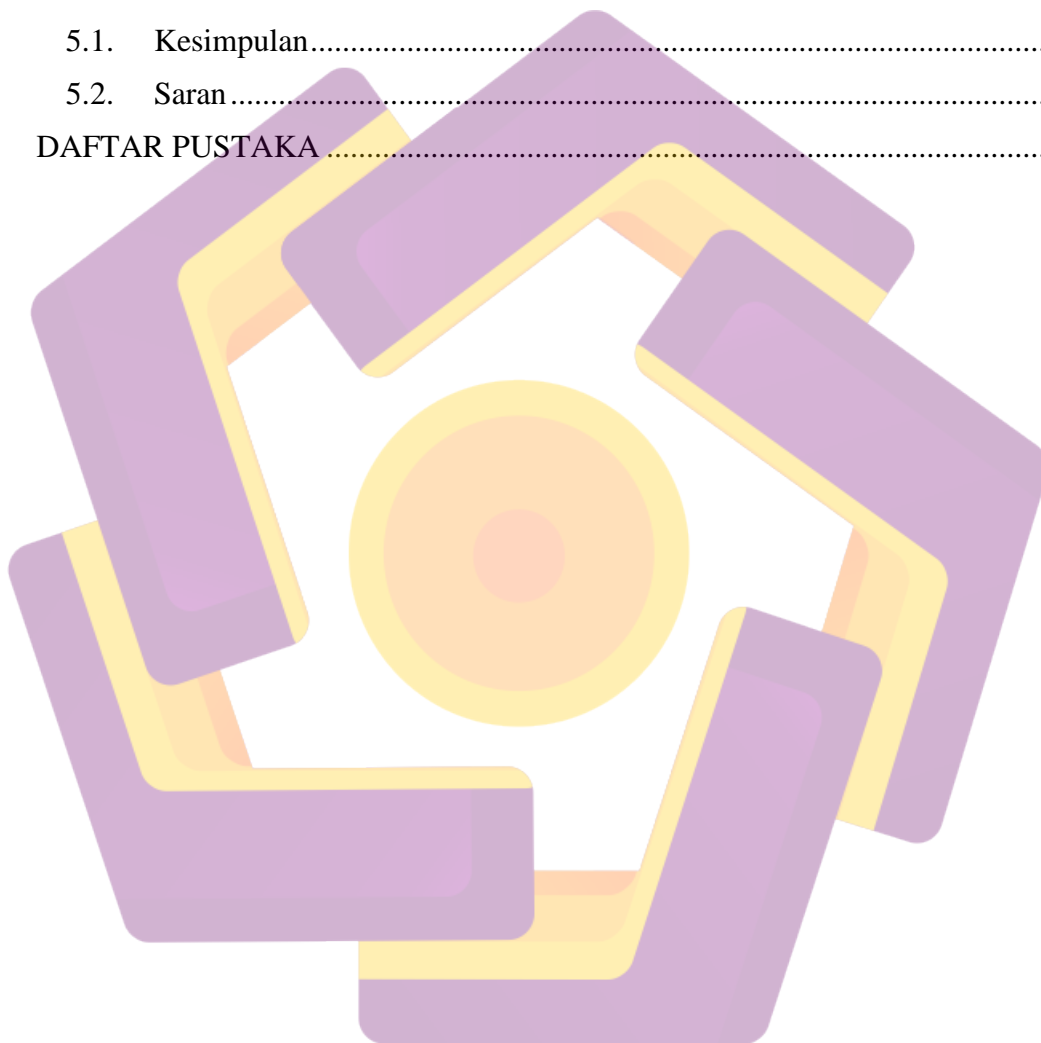
DAFTAR ISI

JUDUL	i
PERSETUJUAN	ii
PENGESAHAN	iii
PERNYATAAN.....	iv
MOTTO	v
PERSEMBAHAN.....	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xv
INTISARI.....	xvi
ABSTRAK.....	xvii
BAB 1	1
PENDAHULUAN	1
1.1. Latar belakang	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah.....	3
1.4. Maksud dan Tujuan Penelitian	3
1.5. Metode Penelitian.....	3
1.5.1. Metode Pengumpulan Data.....	4
1.5.2. Metode Analisis	4

1.5.3. Metode Evaluasi.....	6
1.6. Sistematika Penulisan.....	6
BAB II.....	8
LANDASAN TEORI.....	8
2.1. Tinjauan Pustaka	8
2.2. Data Mining.....	13
2.3. Text Mining	14
2.4. Sentimen Analisis.....	14
2.5. Preprocessing.....	15
2.5.1. Cleansing.....	15
2.5.2. Transform Case	15
2.5.3. Tokenizing.....	15
2.5.4. Normalisasi	16
2.5.5. Stopword Removing	16
2.5.6. Stemming	16
2.6. Data Crawling.....	16
2.7. TF-IDF.....	17
2.8. Confusion Matrix	18
2.9. Multinomial Naïve Bayes.....	19
2.10. Twitter	19
2.10.1. API Twitter	20
2.11. Python.....	20
BAB III	21
METODE PENELITIAN.....	21
3.1. Tinjauan Umum.....	21
3.2. Alat dan Bahan Penelitian	22
3.2.1. Hardware	22
3.2.2. Software	22
3.3. Alur penelitian	22
3.3.1. Crawling data twitter.....	24
3.3.2. Prepossessing	25

3.3.2.1.	Cleansing	25
3.3.2.2.	Transform Case.....	26
3.3.2.3.	Tokenizing	26
3.3.2.4.	Normalisasi	27
3.3.2.5.	Stopword Removing	28
3.3.2.6.	Stemming.....	28
3.3.3.	Pembagian Dataset	29
3.3.4.	TF-IDF	29
3.3.5.	Proses klasifikasi Multinomial Naïve Bayes	33
3.3.5.1.	Penghitungan Kelas Priors.....	33
3.3.5.2.	Penghitungan Probabilitas Kelas Sentimen.....	34
3.3.5.3.	Menentukan Kelas Sentimen	35
3.3.6.	Confusion Matrix	36
3.3.6.1.	Penghitungan Precision.....	37
3.3.6.2.	Penghitungan Recall	37
3.3.6.3.	Penghitungan Accuracy	38
3.3.6.4.	Penghitungan F1	38
BAB IV	39
HASIL DAN PEMBAHASAN	39
4.1.	Prepossessing.....	39
4.1.1.	Cleansing dan Transform case	39
4.1.2.	Tokenizing.....	40
4.1.3.	Normalisasi	41
4.1.4.	Stopword Removal.....	41
4.1.5.	Stemming	42
4.2.	TF-IDF.....	42
4.3.	Pembagian Dataset	42
4.4.	Klasifikasi Multinomial Naïve Bayes	43
4.5.	Confusion Matrix	43
4.6.	Pembuatan Antarmuka	44
4.7.	Pengujian Multinomial Naïve Bayes.....	45
4.7.1.	Pengujian Pertama.....	45

4.7.2. Pengujian Kedua	46
4.7.3. Pengujian Ketiga	47
4.7.4. Pengujian Keempat	48
4.7.5. Pengujian Kelima	49
BAB V.....	51
KESIMPULAN DAN SARAN.....	51
5.1. Kesimpulan.....	51
5.2. Saran.....	51
DAFTAR PUSTAKA.....	52



DAFTAR TABEL

Tabel 2. 1 Tinjauan Pustaka	10
Tabel 2. 2 Confusion Matrix	18
Tabel 3. 1 Spesifikasi Perangkat	22
Tabel 3. 2 Hasil Proses Crawling	24
Tabel 3. 3 Hasil Proses Cleansing	25
Tabel 3. 4 Hasil Proses Transform Case	26
Tabel 3. 5 Hasil Proses Tokenizing	26
Tabel 3. 6 Hasil Proses Normalisasi	27
Tabel 3. 7 Hasil Stopword Removing	28
Tabel 3. 8 Hasil Proses Stemming	29
Tabel 3. 9 Penghitungan Manual TF	30
Tabel 3. 10 TF Positif	31
Tabel 3. 11 TF Netral	32
Tabel 3. 12 TF Negatif	32
Tabel 3. 13 Data Uji	34
Tabel 3. 14 Tabel Hasil Confusion Matrix	37
Tabel 4. 1 Pembagian Dataset	45
Tabel 4. 2 Hasil Confusion Matrix Pengujian Pertama	46
Tabel 4. 3 Hasil pengujian pertama	46
Tabel 4. 4 Hasil Confusion Matrix Pengujian Kedua	47
Tabel 4. 5 Hasil Pengujian Kedua	47

Tabel 4. 6 Hasil Confusion Matrix Pengujian Ketiga 48

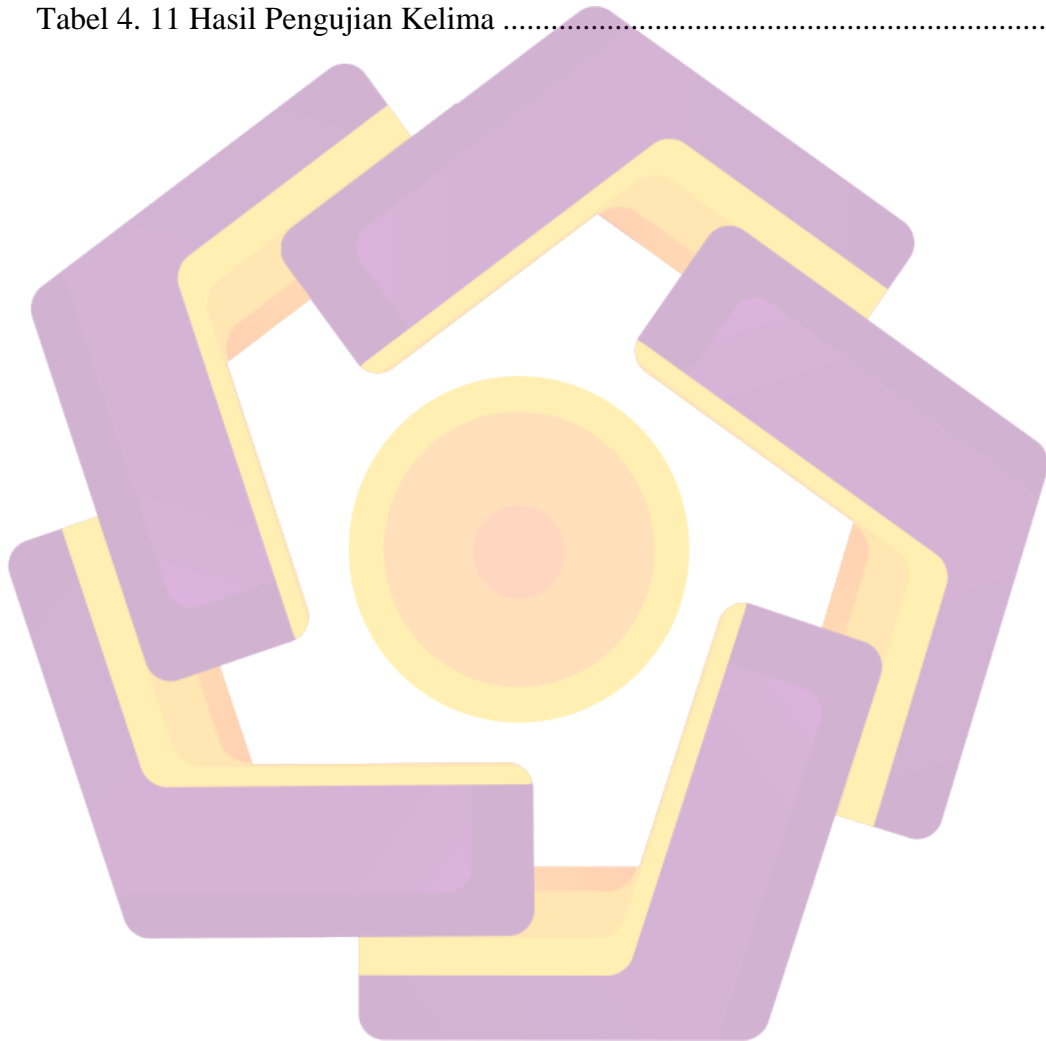
Tabel 4. 7 Hasil Pengujian Ketiga 48

Tabel 4. 8 Hasil Confusion Matrix Pengujian Keempat 49

Tabel 4. 9 Hasil Pengujian Keempat..... 49

Tabel 4. 10 Hasil Confusion Matrix Pengujian Kelima..... 50

Tabel 4. 11 Hasil Pengujian Kelima 50



DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian.....	23
Gambar 4. 1 Implementasi Cleansing	40
Gambar 4. 2 Implementasi Tokenizing.....	40
Gambar 4. 3 Implementasi Normalisasi.....	41
Gambar 4. 4 Implementasi Stopword	41
Gambar 4. 5 Implementasi Stemming.....	42
Gambar 4. 6 Implementasi TF-IDF.....	42
Gambar 4. 7 Implementasi Pembagian Dataset	43
Gambar 4. 8 Implementasi Klasifikasi Multinomial Naive bayes	43
Gambar 4. 9 Implementasi Confusion Matrix	44
Gambar 4. 10 Tampilan Antarmuka.....	44

INTISARI

Industry 4.0 merupakan industri yang berkaitan erat dengan perkembangan teknologi, salah satu teknologi yang berkembang dengan pesat adalah teknologi informasi, dengan media sosial sebagai salah satu contohnya, Twitter adalah salah satu media penyebaran informasi, mulai dari berita, kritik, opini dengan data yang berlimpah. Sehingga memudahkan untuk melakukan Analisa terhadap suatu produk atau jasa. Penulis tertarik untuk melakukan penelitian mengenai sentimen analisis terhadap provider by.u menggunakan metode multinomial naïve bayes.

Penelitian ini menggunakan dataset yang berasal dari twitter dengan jumlah data yang digunakan berjumlah 1050 data. Dari data yang didapat akan dilakukan preprocessing terlebih dahulu agar data siap digunakan untuk mengetahui apakah metode multinomial naïve bayes dapat digunakan untuk melakukan sentiment analisis.

Berdasarkan hasil penelitian yang telah dilakukan tingkat akurasi yang dihasilkan metode Multinomial naïve bayes mencapai 79% dengan perbandingan dataset 840 data latih dan 210 data uji. Dari hasil akurasi tersebut Multinomial naïve bayes dapat digunakan dalam melakukan sentimen analisis dengan hasil yang cukup baik.

Kata Kunci: Sentimen analisis, Multinomial Naïve Bayes, Twitter, By.U.

ABSTRAK

Industry 4.0 is an industry that is closely related to technological developments, one of the fastest growing technologies is information technology, with social media as an example, Twitter is one of the media for disseminating information, ranging from news, criticism, opinions with abundant data. Making it easier to carry out an analysis of a product or service. The author is interested in conducting research on sentiment analysis of provider By.u using the multinomial naïve Bayes method.

This study uses a dataset derived from Twitter with the amount of data used is 1050 data. From the data obtained, preprocessing will be carried out first so that the data is ready for use to determine whether the naïve Bayes multinomial method can be used to perform sentiment analysis.

Based on the results of research that has been carried out, the level of accuracy produced by the multinomial naïve Bayes method reaches 79% with a comparison of 840 training data and 210 test data. From the results of this accuracy, Naive Bayes Multinomial can be used in conducting sentiment analysis with quite good results.

Keyword : Sentiment Analysis, Multinomial Naïve Bayes, Twitter, By.U.