

**ANALISIS SENTIMEN PRODUK HP TERBARU DARI BRAND ITEL  
MENGUNAKAN NAIVE BAYES CALSSIFIER**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi *Informatika*



disusun oleh

**ABI DAFFA ARILA**

**21.21.1538**

Kepada

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

**2022**

**ANALISIS SENTIMEN PRODUK HP TERBARU DARI BRAND ITEL  
MENGUNAKAN NAIVE BAYES CALSSIFIER**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana Program  
*Studi Informatika*



disusun oleh

**ABI DAFFA ARIILA**

**21.21.1538**

Kepada

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

**2022**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS SENTIMEN PRODUK HP TERBARU DARI BRAND ITEL  
MENGUNAKAN NAIVE BAYES CALSSIFIER**

yang disusun dan diajukan oleh

Abi Daffa Arifa

21.21.1538

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 22 Juli 2023

Dosen Pembimbing,

Kusrini, Prof., Dr., M.Kom

NIK. 190302106

HALAMAN PENGESAHAN  
SKRIPSI  
ANALISIS SENTIMEN PRODUK HP TERBARU DARI BRAND ITEL  
MENGUNAKAN NAIVE BAYES CALSSIFIER

yang disusun dan diajukan oleh

Abi Daffa Arila

21.21.1538

Telah dipertahankan di depan Dewan Penguji  
pada tanggal <tanggal ujian>

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Mei P Kurniawan, M.Kom  
NIK. 190302187

Yuli Astuti, M.Kom  
NIK. 190302146

Kusrini, Prof., Dr., M.Kom  
NIK. 190302106



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 20 Juli 2023

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom.  
NIK. 190302096

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

**Nama mahasiswa : Abi Daffa Arila**

**NIM : 21.21.1538**

Menyatakan bahwa Skripsi dengan judul berikut:

**Analisis Sentimen Produk Hp Terbaru dari Brand ITEL menggunakan Naïve Bayes Classifier**

**Dosen Pembimbing : Prof. Kusriani, Dr., M.Kom**

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 20 Juli 2023

Yang Menyatakan,



Abi Daffa Arila

## KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kepada Allah Subhanahu Wa Ta'ala yang telah memberikan limpahan rahmat, hidayah, serta kemudahan-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis Sentimen Produk Hp Terbaru dari Brand ITEL menggunakan Naïve Bayes Classifier" dengan segenap usaha dan semampu yang penulis bisa. Alhamdulillah, skripsi ini telah disusun sebagai salah satu syarat untuk mendapatkan gelar Sarjana pada program Informatika Ilmu Komputer Universitas Amikom Yogyakarta.

Dalam penulisan skripsi ini, tentu penulis tidak lepas dari dukungan, motivasi, bimbingan, dan saran dari berbagai pihak yang telah banyak membantu dalam proses pembuatan skrip ini, baik secara lisan maupun secara tertulis. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih kepada :

1. Allah Subhanahu Wa Ta'ala atas segala limpahan nikmat-Nya, rasa sabar dan dimampukan, sehingga penulis mampu menyelesaikan laporan tugas akhir ini dengan baik dan lancar.
2. Bapak Ari dan Ibu Mila, selaku orang tua penulis yang selalu memberikan dukungan baik lahiriah maupun batiniah, semangat, nasihat, dan segala aspek baik secara moril maupun materil.
3. Azhom, Atha, Aura dan Alyaa selaku saudara/i penulis yang selalu memberikan dukungan serta semangat kepada penulis.
4. Prof Kusri. Dr., M.Kom selaku dosen pembimbing skripsi yang hingga saat ini selalu memberikan arahan, koreksi, bimbingan, dan tentu semangat.
5. Seluruh staf dan pengajar program studi S1 Informatika atas bantuan dan dukungan yang diberikan kepada penulis.
6. Uci Agustina, Redho Kurniawan, Ray Hari, Diah Ayu, Mas Fuad, Suci Share Putri selaku teman yang selalu memberikan dukungan dan motivasi

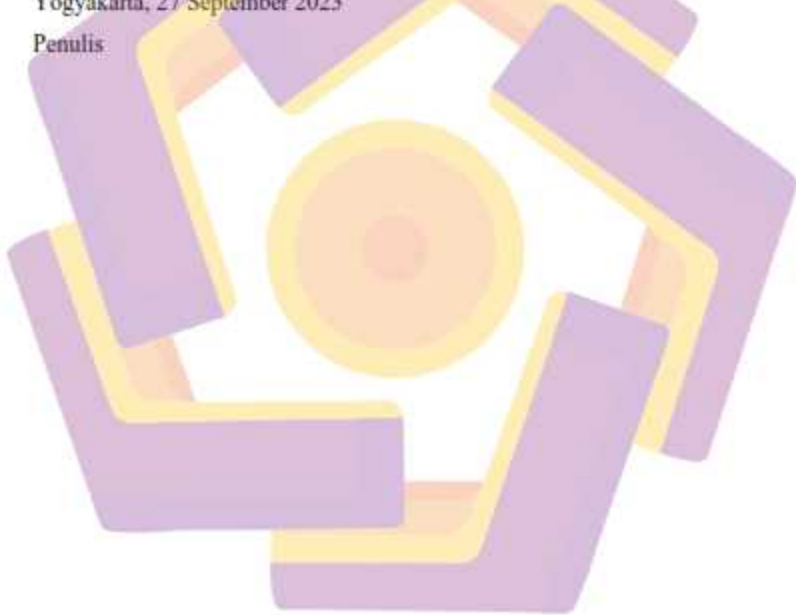


untuk terus berproses dalam menyelesaikan skripsi ini.

7. Bapak dan Ibu dosen serta teman-teman di Silvanus yang terus memberikan dorongan dan semangat untuk terus berusaha menggapai apa yang diinginkan, sekaligus menularkan kebiasaan yang baik dan bermanfaat untuk terus bertumbuh.
8. Teman-teman yang tidak dapat sebutkan satu persatu yang dengan senyuman disetiap canda dan tawa, do'a disetiap kebersamaan, dan sumbangsi semangat disetiap langkah yang membersamai penulis hingga saat ini

Yogyakarta, 27 September 2023

Penulis



## DAFTAR ISI

HALAMAN JUDUL .....	1
DAFTAR ISI.....	i
DAFTAR TABEL.....	iv
DAFTAR GAMBAR.....	v
DAFTAR LAMPIRAN.....	vii
INTISARI.....	viii
ABSTRACT.....	ix
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penulisan.....	4
<b>BAB II TINJAUAN PUSTAKA</b> .....	<b>5</b>
2.1 Studi Literatur.....	5
2.2 Dasar Teori.....	12
2.2.1. Analisis Sentimen.....	12
2.2.2. Youtube.....	12
2.2.3. Vader Lexicon.....	13
2.2.4. Scrapping.....	13
2.2.5. Text Mining.....	14
2.2.6. Text Preprocessing.....	14



2.2.7. Term Frequency-Inverse Document Frequency Klasifikasi .....	16
2.2.8. Machine Learning .....	17
2.2.9. Naive Bayes Classifier .....	18
2.2.10. Confusion Matrix .....	19
<b>BAB III METODE PENELITIAN .....</b>	<b>21</b>
3.1 Objek Penelitian .....	21
3.2 Alur Penelitian .....	21
3.2.1 Scrapping Data .....	21
3.2.2 Labeling Dataset .....	22
3.2.3 Pembagian Dataset .....	22
3.3 Alat dan Bahan .....	22
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>24</b>
4.1 Scrapping Dataset .....	24
4.2 Preprocessing Text .....	28
4.2.1 Clean Text .....	28
4.2.2 Casefolding .....	28
4.2.3 Tokenizing .....	29
4.2.4 Stopwords .....	30
4.2.5 Lemmatization .....	30
4.2.6 De Tokenize .....	31
4.2.7 Count Words .....	32
4.2.8 Drop Rows .....	32
4.3 Labelling Dataset .....	33
4.4 Splitting Dataset .....	36

4.5 Naive Bayes Classifier.....	36
4.6 Pengujian Pada Perbandingan Data Training 70:30 .....	38
4.6.1 Imbalance & Default Parameter.....	38
4.6.2 Imbalance & Hyperparameter Tuning .....	38
4.6.3 Balance & Default Parameter .....	39
4.6.4 Balance & Hyperparameter Tuning.....	39
4.7 Pengujian Pada Perbandingan Data Training 80:20 .....	40
4.7.1 Imbalance & Default Parameter.....	40
4.7.2 Imbalance & Hyperparameter Tuning .....	40
4.7.3 Balance & Default Parameter .....	40
4.7.4 Balance & Hyperparameter Tuning .....	41
4.8 Visualisasi Kemunculan Kata Setiap Sentimen .....	41
4.9 Diskusi Hasil.....	43
4.9.1 Perbandingan Hasil .....	43
4.9.2 Metode dan Perlakuan Terbaik .....	44
<b>BAB V PENUTUP .....</b>	<b>46</b>
5.1 Kesimpulan .....	46
5.2 Saran .....	47
<b>REFERENSI .....</b>	<b>48</b>
<b>LAMPIRAN.....</b>	<b>52</b>

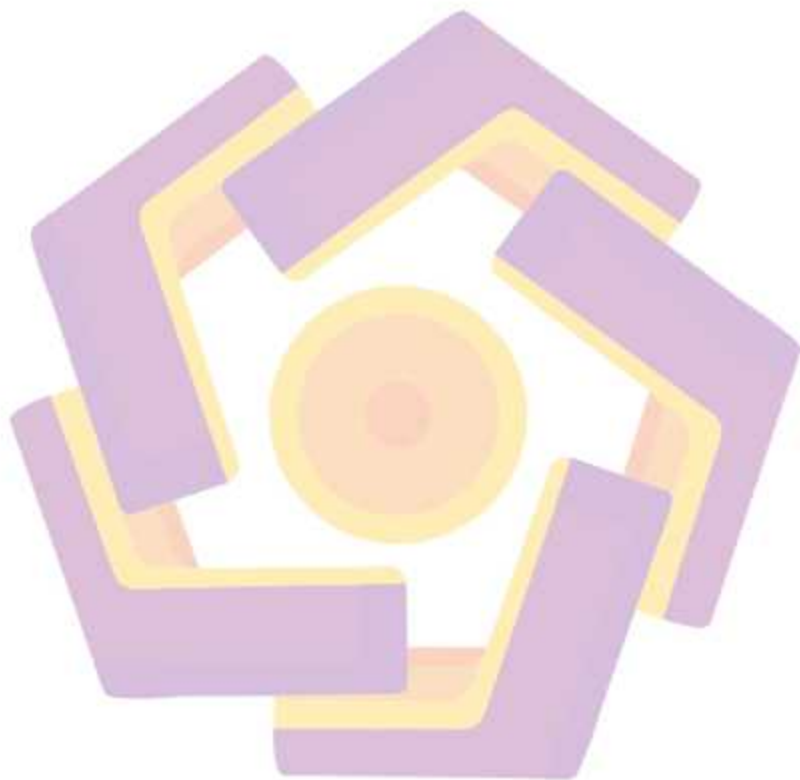
## DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian .....	7
Tabel 2.1 Confusion Matrix .....	19
Tabel 2.2 Rumus Perhitungan Confusion Matrix .....	20
Tabel 4.2 Hasil Clean Text dan Casefolding .....	29
Tabel 4.3 Hasil Tokenizing .....	29
Tabel 4.4 Hasil Stopwords .....	30
Tabel 4.5 Hasil Lemmatization .....	31
Tabel 4.5 Hasil De Tokenize .....	31
Tabel 4.6 Hasil Count Words .....	32
Tabel 4.7 Hasil Translation .....	33
Tabel 4.8 Hasil Vader Lexiocn .....	34
Tabel 4.8 Parameter Optimal MultinomialNB .....	44

## DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian .....	21
Gambar 3.2 Alur Scrapping Data.....	21
Gambar 3.6 Alur Labeling Dataset .....	22
Gambar 4.1 Tampilan Google Developer Console.....	24
Gambar 4.2 Halaman <i>New Project</i> .....	25
Gambar 4.3 Layanan API Youtube V3.....	25
Gambar 4.4 Halaman <i>Create Credentials</i> .....	26
Gambar 4.5 Kode API Key .....	26
Gambar 4.6 Import Library Scrapping.....	26
Gambar 4.7 Kode Ekstraksi Komentar YouTube .....	27
Gambar 4.8 Kode Integrasi API YouTube .....	27
Gambar 4.9 Kode Export Dataset .....	27
Gambar 4.10 Kode Clean Text .....	28
Gambar 4.11 Kode Casefolding.....	28
Gambar 4.12 Kode Tokenizing .....	29
Gambar 4.13 Kode Stopwords .....	30
Gambar 4.14 Kode Lemmatization.....	30
Gambar 4.15 Kode De Tokenize.....	31
Gambar 4.15 Kode Count Words.....	32
Gambar 4.16 Kode Drop Rows.....	32
Gambar 4.17 Hasil Drop Rows .....	33
Gambar 4.18 Kode Translation .....	33
Gambar 4.19 Kode Vader Lexicon .....	34

Gambar 4.20 Hasil <i>Labeling Dataset</i> .....	35
Gambar 4.21 Pelabelan Vader Lexicon .....	35
Gambar 4.21 Kode <i>Splitting Dataset</i> .....	36
Gambar 4.22 Hasil Test Non-Balancing .....	38



## DAFTAR LAMPIRAN

Gambar 1. Scrapping Data.....	52
Gambar 2. Import Dataset.....	53
Gambar 3. Library Setup.....	53
Gambar 4. Data Review.....	53
Gambar 5. Clean Text.....	53
Gambar 6. Casefolding.....	54
Gambar 7. De Acronim Text.....	55
Gambar 8. Tokenizing.....	55
Gambar 9. Stopwords.....	55
Gambar 10. Lemmatization.....	56
Gambar 11. De Tokenize.....	56
Gambar 12. Count Words.....	56
Gambar 13. Drop Rows.....	56
Gambar 14. Translate Dataset.....	56
Gambar 15. Labeling Sentiment Dataset.....	57
Gambar 16. TF-IDF.....	57
Gambar 17. Splitting Dataset.....	57
Gambar 18. Balancing Dataset.....	57
Gambar 19. Model Train Set.....	58
Gambar 21. Confusion Matrix.....	58
Gambar 22. Visualisasi Confusion Matrix.....	58



## INTISARI

Produk terbaru dari brand ITEL menjadi viral beberapa hari setelah diluncurkan karena menawarkan spesifikasi yang bagus dengan harga yang terjangkau. Penelitian dilakukan untuk memahami sentimen konsumen terhadap produk tersebut. Data yang digunakan dalam penelitian ini adalah komentar-komentar pada video review produk HP ITEL di channel Gadgetin dengan judul: "Ribuan orang minta saya buat bahas HP 1,3 juta ini". Pelabelan menggunakan *Vader Lexicon* menghasilkan 1970 data bersih dari total 2000 data, dengan 1125 sentimen netral, 656 sentimen positif, dan 205 sentimen negatif. Penelitian ini menggunakan Metode Naive Bayes Classifier untuk mengolah hasil konversi data dari TF-IDF. Dilakukan analisis terhadap pola sentimen dan pengaruhnya terhadap persepsi konsumen terhadap produk HP ITEL. Hasil analisis menggunakan Multinomial Naive Bayes (MultinomialNB) dengan parameter default, pembagian dataset sebesar 70:30, dan TF-IDF dengan parameter *ngram\_range* (1,2) menunjukkan tingkat akurasi sebesar 70% setelah dilakukan *balancing* pada dataset karena ketidakseimbangan pembagian dataset. Kesimpulan yang diperoleh adalah sebagian besar masyarakat Indonesia cenderung memiliki sentimen biasa saja (netral) terhadap produk handphone terbaru dari brand ITEL, namun sebagian besar komentar masyarakat bersifat jelek (negatif) dikarenakan konotasi negatif yang terkait dengan nama brand, terutama pada daerah Jawa. Temuan penelitian ini diharapkan dapat memberikan wawasan yang berharga bagi brand ITEL dalam memahami persepsi dan tanggapan konsumen, serta membantu dalam meningkatkan kualitas produk, pengalaman pengguna, dan pengembangan strategi pemasaran yang lebih efektif.

**Kata kunci:** Analisis Sentimen, HP, ITEL, YouTube, Naive Bayes Classifier, Persepsi Konsumen.

## ABSTRACT

The latest product from the Itel brand went viral a few days after its launch due to its affordable price and impressive specifications. The research was conducted to understand consumer sentiment towards the product. The data used in this study consisted of comments on the Itel HP product review video on the Gadgetin channel with the title: "Thousands of people asked me to review this 1.3 million rupiah phone." Labeling using the Vader Lexicon resulted in 1,970 clean data out of a total of 2,000, with 1,125 neutral sentiments, 656 positive sentiments, and 205 negative sentiments. This research utilized the Naïve Bayes Classifier method to process the converted data using TF-IDF. An analysis was performed on sentiment patterns and their impact on consumer perceptions of the Itel HP product. The analysis results using Multinomial Naive Bayes (MultinomialNB) with default parameters, a dataset split of 70:30, and TF-IDF with ngram\_range (1,2) parameter showed an accuracy rate of 70% after balancing the dataset due to the imbalance in dataset distribution. The conclusion drawn was that the majority of Indonesian society tends to have a neutral sentiment towards the latest Itel brand mobile phone product. However, a significant portion of the public's comments were negative due to negative connotations associated with the brand name, especially in Java. The findings of this research are expected to provide valuable insights for the Itel brand in understanding consumer perceptions and responses, as well as aiding in enhancing product quality, user experience, and the development of more effective marketing strategies.

**Keyword:** Sentiment Analysis, HP, Itel, YouTube, Naive Bayes Classifier, Consumer Perceptions.