HALAMAN MOTTO

“Hidup bukan saling mendahului, Bermimpilah sendiri-sendiri”

-Besok Mungkin Kita Sampai (Hindia)

“Tidak ada mimpi yang terlalu tinggi dan tidak ada mimpi yang patut di remehkan. Lambungkan setinggi yang kau inginkan dan gapailah dengan selayaknya yang kau harapkan.”

-Maudy Ayunda

“Tidak ada kata terlambat selagi kamu mau berusaha dengan sungguh-sungguh”

-Widya Kusuma Whardani

HALAMAN PERSEMBAHAN

Puji syukur penulis penulis panjatkan kepada Allah SWT yang telah memberikan keikhlasan, kekuatan dan kesabaran sehingga Skripsi ini dapat terselesaikan.

Skipsi ini penulis persembahkan kepada:

1. “Widya Kusuma Whardani” ya terimakasi kepada diri saya sendiri, karena telah bertahan dan berjuang hingga sampai di titik ini. Terima kasih karena tetap kuat menghadapi berbagai proses kehidupan yang tidak selalu berjalan sesuai harapan yang tidak banyak orang tau, serta tidak menyerah meskipun ada banyak hal yang harus dihadapi. atas segala kerja keras, doa, dan pengorbanan hingga akhirnya mampu menyandang gelar S.Kom. Menjadi anak perempuan pertama tidaklah mudah, tetapi semua proses yang telah dilalui berhasil membawa langkah ini sampai pada tujuan yang diimpikan. Semoga selalu memiliki cara untuk membahagiakan diri sendiri, terus bertumbuh menjadi pribadi yang lebih baik, dan tidak pernah lupa menghargai setiap proses yang telah dilewati.
2. Bapak Ir. R. Muhamad Ikhsanuddin, S. Kom., M.Cs. selaku dosen pembimbing, serta Bpak/Ibu penguji, atas kesabaran, arahan, dan waktu yang diberikan selama proses bimbingan, Ilmu dan masukan yang diberikan menjadi bagian penting dari terselesai kan nya skripsi ini.
3. Kedua orang tua penulis yang telah menjadi bagian penting dalam hidup ini. Terima kasih atas doa, restu, dukungan, serta pengorbanan yang diberikan telah menjadi kekuatan terbesar dalam setiap langkah perjalanan hidup ini.

4. Untuk ke-3 adikku Faran Kusuma, Syariisah Zahira Amir, dan adik terakhir ku Ghany Yazid Khairi, alasan penulis untuk terus berjuang dan tidak menyerah. Semoga karya kecil ini menjadi bukti bahwa kerja keras dan doa tidak pernah sia-sia. Penulis berharap akan selalu percaya bahwa tidak ada mimpi yang terlalu tinggi untuk diraih.
5. Saudara-saudara yang selalu memberikan semangat, doa, dan dukungan ketika penulis merasa lelah dan hampir menyerah. Kehadiran kalian adalah penguat yang membuat penulis mampu melangkah hingga sejauh ini.
6. Sahabat penulis Fitroh Widi Astuti, yang tetap menjadi bagian dari perjalanan hidup penulis sejak masa SMP hingga perkuliahan, selalu menemani dan mendengarkan setiap cerita dalam suka maupun duka, serta tetap sama meski kesibukan dan teman baru datang silih berganti.
7. Terima kasih kepada ibu-ibu rewang yang telah menjadi bagian dari keluarga kedua penulis, khususnya Bu Tri, Mba Rini, Mba Yuli, Bu Tuti, Mba Septi, Lik Siti, serta ibu-ibu lainnya yang tidak dapat penulis sebutkan satu per satu. Terima kasih atas doa, dukungan, dan kehangatan yang selalu diberikan di setiap proses penulis.
8. Dan tidak lupa teman-teman penulis yang tidak dapat disebutkan satu per satu, teman seperjuangan khususnya IKRC dari semester 1 dan teman-teman IKRB yang sudah meng izinkan untuk bisa bergabung di beberapa matkul dan pindah kembali masuk kelas IKRA karena Ilkom terbagi menjadi 2 kelas, yang telah menjadi bagian dari perjalanan panjang ini. Terimakasih atas kebersamaan, tawa, doa, dan dukungan yang telah

diberikan. Kehadiran kalian membuat perjalanan ini penuh warna dan bermakna.

ABSTRAK

Komentar pada platform YouTube memuat opini masyarakat dalam jumlah besar sehingga sulit dianalisis secara manual. Penelitian ini bertujuan mengklasifikasikan sentimen komentar YouTube terhadap iPhone 16 ke dalam kategori positif, negatif, dan netral menggunakan algoritma *Naive Bayes*, serta mengukur kinerjanya. Data berupa 4.199 komentar dari tiga video ulasan iPhone 16 yang dikumpulkan melalui *web scraping* dengan Apify. Tahapan penelitian meliputi *cleaning*, *case folding*, *tokenizing*, *stopword removal*, pelabelan otomatis berbasis kamus (*lexicon-based*), pembobotan TF-IDF, pembagian data latih dan data uji dengan rasio 70:30, dan klasifikasi menggunakan *Naive Bayes* pada RapidMiner. Hasil pelabelan menunjukkan sentimen negatif paling dominan dengan 1.806 komentar (43,01%), diikuti netral 1.324 komentar (31,53%) dan positif 1.069 komentar (25,46%). Evaluasi pada 1.260 data uji menghasilkan *accuracy* 49,13%, rata-rata *recall* 40,26%, rata-rata *precision* 81,94%, dan rata-rata *F1-score* 33,39%. Model cenderung memprediksi komentar sebagai negatif (*recall* kelas negatif 100%, netral 13,60%, positif 7,17%). Hal ini diduga dipengaruhi oleh ketidakseimbangan data antarkelas, pelabelan berbasis kamus yang tidak menangani negasi dan sarkasme, serta penggunaan bahasa tidak baku. Dengan demikian, kinerja *Naive Bayes* pada penelitian ini masih tergolong rendah dan perlu ditingkatkan melalui penyeimbangan data, perbaikan pelabelan, dan perbandingan dengan algoritma lain.

Kata Kunci: Analisis Sentimen, iPhone 16, YouTube, *Naive Bayes*, TF-IDF.

ABSTRACT

YouTube comments contain a large volume of public opinion that is difficult to analyze manually. This study aims to classify YouTube comment sentiment toward the iPhone 16 into positive, negative, and neutral categories using the Naive Bayes algorithm and to measure its performance. The data consist of 4,199 comments from three iPhone 16 review videos, collected through web scraping with Apify. The research stages include cleaning, case folding, tokenizing, stopword removal, automatic lexicon-based labeling, TF-IDF weighting, a 70:30 split of training and testing data, and Naive Bayes classification in RapidMiner. The labeling results show that negative sentiment is dominant with 1,806 comments (43.01%), followed by neutral with 1,324 comments (31.53%) and positive with 1,069 comments (25.46%). Evaluation on 1,260 test samples yielded an accuracy of 49.13%, an average recall of 40.26%, an average precision of 81.94%, and an average F1-score of 33.39%. The model tends to predict comments as negative (recall of 100% for negative, 13.60% for neutral, and 7.17% for positive). This is presumably influenced by class imbalance, lexicon-based labeling that does not handle negation and sarcasm, and the use of non-standard language. Therefore, the performance of Naive Bayes in this study remains low and needs improvement through data balancing, better labeling, and comparison with other algorithms.

Keywords: *Sentiment Analysis, iPhone 16, YouTube, Naive Bayes, TF-IDF.*

KATA PENGANTAR

Puji syukur ke hadirat Allah SWT atas Rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan tugas akhir berupa skripsi ini yang berjudul "**SENTIMEN ANALISIS PRODUK IPHONE 16 PADA KOMENTAR YOUTUBE MENGGUNAKAN METODE *NAIVE BAYES***" dengan tepat waktu. Skripsi ini ditulis sebagai salah satu syarat untuk mendapatkan gelar Sarjana Strata 1 (S1) pada program studi Ilmu Komputer Universitas Putra Bangsa Kebumen.

Skripsi ini menjadi salah satu bentuk pencapaian penulis dalam mengaplikasikan ilmu yang diperoleh selama masa perkuliahan ke dalam praktik nyata, khususnya dalam kegiatan penelitian ilmiah. Proses penyusunan skripsi ini dilakukan dengan segala kemampuan yang dimiliki penulis, memadukan teori-teori yang relevan serta dukungan dari berbagai literatur yang menunjang. Penulis menyadari sepenuhnya bahwa karya ini masih jauh dari sempurna mengingat keterbatasan ilmu dan pengalaman yang dimiliki.

Pada kesempatan yang berbahagia ini, penulis ingin menyampaikan rasa terima kasih yang sedalam-dalamnya kepada semua pihak yang telah memberikan dukunga, baik dalam bentuk bantuan nyata, dorongan semangat, maupun motivasi yang tidak ternilai harganya. Oleh karena itu, dengan penuh rasa hormat dan penghargaan, penulis menyampaikan terima kasih kepada:

1. Bapak Dr. Gunarso Wiwoho, S.E., M.M, selaku Rektor Universitas Putra Bangsa

2. Bapak Yulianto, S. Kom., M. Kom., selaku Dekan Fakultas Sains dan Teknologi
3. Bapak Awaludin A'bid, S. Kom., M. Kom., selaku Kepala Program Studi Ilmu Komputer Universitas Putra Bangsa, atas segala dukungan, arahan, serta fasilitas yang diberikan kepada penulis selama masa perkuliahan.
4. Ir. Rohmatulloh Muhamad Ikhsanuddin, S. Kom, M. Cs., selaku dosen pembimbing yang dengan penuh kesabaran, keikhlasan, serta ketelatenan telah membimbing, memberikan arahan, dan selalu memotivasi penulis hingga skripsi ini dapat terselesaikan.
5. Seluruh dosen dan karyawan Universitas Putra Bangsa, yang telah memberikan ilmu pengetahuan, pengalaman, serta bantuan dalam berbagai bentuk selama penulis menempuh pendidikan hingga terselesaikannya skripsi ini.

Penulis menyadari sepenuhnya bahwa skripsi ini masih jauh dari sempurna. Besar harapan penulis, semoga karya ilmiah ini dapat memberikan manfaat, baik bagi penulis sendiri maupun bagi pembaca, serta dapat menjadi bagian kecil dari pengembangan ilmu pengetahuan.

Kebumen, 6 September 2026
Penulis,

Widya Kusuma Whardani

DAFTAR ISI

HALAMAN JUDUL i
HALAMAN PERSETUJUAN iii
HALAMAN PENGESAHAN (hasil scan) iv
HALAMAN BEBAS PLAGIARISME v
HALAMAN MOTTO vi
HALAMAN PERSEMBAHAN vii
ABSTRAK x
ABSTRACT xi
KATA PENGANTAR xii
DAFTAR ISI xiv
DAFTAR TABEL xvi
DAFTAR GAMBAR xvii
DAFTAR LAMPIRAN xviii
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Penelitian 9
1.3 Tujuan Penelitian 9
1.4 Batasan Penelitian 10
1.5 Manfaat Penelitian 11
BAB II KAJIAN PUSTAKA 13
2.1 Tinjauan Teori 13

2.1.1 *Platform* Youtube 13
2.1.2 *Text Mining* 16
2.1.3 Analisis Sentimen 18
2.1.4 Metode *Naive Bayes* 20
2.1.5 *Rapid Miner* 26
2.1.6 *Web Scraping* 26
2.2 Penelitian Terdahulu 29
2.3 Kerangka Pemikiran 31
2.4 Keaslian Penelitian 33
BAB III 35
METODOLOGI PENELITIAN 35
3.1 Tahapan Penelitian 35
3.2 Metode Penelitian 39
3.3 Metode Pengumpulan Data 40
3.4 Metode Analisis Data 43
3.5 Alat dan Bahan 51
3.6 Prosedur Pengujian 54
Akhirnya SAH LEGAL! Review iPhone 16 & 16 Pro Indonesia! 55
REVIEW iPhone 16 Indonesia 55
https://youtu.be/7cOr8WPPq9w?si=8rh6I3DHjoh5kFmJ 55
BAB IV HASIL DAN PEMBAHASAN 59
4.1 Pengumpulan Data 59
4.1.1 *Scraping Data* 59

4.1.2 Pelabelan Data 61
4.1.3 Distribusi Sentimen..... 63
4.2 Implementasi *Pre-Processing*..... 64
4.2.1 *Cleaning*..... 65
4.2.2 *Case Folding*..... 68
4.2.3 TF-IDF 70
4.2.4 *Stopword Removal* 73
4.2.5 *Tokenize* 75
4.2.6 Algoritma Klasifikasi Dengan *Naive Bayes*..... 76
4.2.7 *Confision Matrix* 79
4.3 Pembagian Data Latih dan data Uji 80
BAB V PENUTUP..... 82
5.1 Kesimpulan 82
5.2 Saran 83
DAFTAR PUSTAKA 85
LAMPIRAN..... 87

DAFTAR TABEL

Tabel II - 1. Penanganan Singkatan dan Bahasa Tidak Baku 18
Tabel II - 2. Penelitian Terdahulu 29
Tabel III - 1. Pelabelan Sentimen..... 50
Tabel III - 2. *Library* Python..... 53
Tabel III - 3. Sumber Video 55
Tabel IV - 1. Kamus *Stopword* 74

DAFTAR GAMBAR

Gambar II - 1. Kerangka Pemikiran 33
Gambar III - 1. Alur Penelitian 35
Gambar III - 2. Metode Analisis Data 45
Gambar III - 3. Contoh Komentar YouTube 57
Gambar IV - 1. Hasil *Scraping Data* 60
Gambar IV - 2. Pelabelan Sentimen 61
Gambar IV - 3. Proses Perlabelan Otomatis 62
Gambar IV - 4. Hasil Perlabelan Otomatis 62
Gambar IV - 5. *Process Cleaning* 1 65
Gambar IV - 6. *Process Cleaning* 2 66
Gambar IV - 7. *Process Cleaning* 3 66
Gambar IV - 8. Hasil *Cleaning* 68
Gambar IV - 9. *Case Folding* 69
Gambar IV - 10. TF - IDF 70
Gambar IV - 11. Proses TF-IDF 71
Gambar IV - 12. Hasil Pengujian *TF IDF* 72
Gambar IV - 13. Hasil *Stopword Removal* 75
Gambar IV - 14. Hasil *Tokenzing* 76
Gambar IV - 15. Model Proses Algoritma *Naive Bayes* 77
Gambar IV - 16. Hasil Pengujian 79
Gambar IV - 17. *Performance Vector* 79
Gambar IV - 18. *Accuracy* 80

DAFTAR LAMPIRAN

Lampiran I. Kartu Bimbingan Skripsi..... 88
Lampiran II. Kartu Seminar Proposal 89
Lampiran III. Dataset Awal 90
Lampiran IV Data Sentimen Labeling 112