

**ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP
RKUHP PADA MEDIA SOSIAL X MENGGUNAKAN NAIVE
BAYES CLASSIFIER
SKRIPSI**

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



disusun oleh
IVANDA ARY SUGITA
19.11.3236

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP
RKUHP PADA MEDIA SOSIAL X MENGGUNAKAN NAIVE
BAYES CLASSIFIER**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi *informatika*



disusun oleh

IVANDA ARY SUGITA

19.11.3236

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP
RKUHP PADA MEDIA SOSIAL X MENGGUNAKAN NAIVE BAYES
CLASSIFIER**

yang disusun dan diajukan oleh

IVANDA ARY SUGITA

19.11.3236

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal <tanggal ujian>

Dosen Pembimbing,



Windha Mega Pradnya Dhuhiya, M.Kom.

NIK. 190302185

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP
RKUHP PADA MEDIA SOSIAL X MENGGUNAKAN NAIVE BAYES
CLASSIFIER

yang disusun dan diajukan oleh

IVANDA ARY SUGITA

19.11.3236

Telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Desember 2024

Susunan Dewan Penguji

Nama Penguji

Majid Rahardi, S.Kom., M.Eng
NIK. 190302393

Surya Tri Atmaja Ramadhani, S.Kom., M.Eng
NIK. 190302481

Windha Mega Pradnya Duhita, S.Kom., M.Kom
NIK. 190302185

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 19 Desember 2024

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : IVANDA ARY SUGITA
NIM : 19.11.3236

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP
RKUHP PADA MEDIA SOSIAL X MENGGUNAKAN NAIVE BAYES
CLASSIFIER**

Dosen Pembimbing: Windha Mega Pradnya Duhita, 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, 19 Desember 2024

Yang Menyatakan,



Ivanda Ary Sugita

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur dan bahagia telah menyelesaikan laporan tugas akhir ini yang tak luput dari doa-doa dan dukungan dari orang-orang tercinta yang selalu memberikan support. Dengan rasa bangga dan syukur saya haturkan rasa syukur dan terimakasih saya kepada:

1. Allah SWT karena hanya atas izin dan karunianya lah skripsi ini dapat dibuat dan selesai pada waktunya.
2. Kedua orang tua saya bapak Surono dan ibu Retno Ningsih , yang telah memberikan dukungan moril maupun materi serta doa yang tiada henti untuk kesuksesan saya, karena tiada kata seindah lantunan doa dan tiada doa yang paling khusyuk selain doa yang terucap dari orang tua.
3. Windha Mega Pradnya Dhuhiya, M.Kom. selaku dosen pembimbing saya yang telah membimbing dalam proses pembuatan skripsi ini.
4. Bapak Hanif Al Fatta, S.Kom., M.Kom selaku Dekan Fakultas Ilmu Komputer.
5. Bapak serta Ibu Dosen prodi Sistem Informasi.
6. Teman-teman prodi Informatika 10 angkatan 2019 yang telah berjuang bersama.
7. Diri saya sendiri yang telah berjuang dan semangat sampai di titik ini.

KATA PENGANTAR

Alhamdulillah puji syukur penulis panjatkan kepada ALLAH SWT atas rahmat dan karunianya yang telah senantiasa membimbing dan memudahkan jalan penulis dalam menyelesaikan penulisan tugas akhir ini. Tak lupa penulis ucapkan terimakasih kepada Dosen Pembimbing penulis, M.Kom dan Tim Dosen Penguji yaitu.... Dan semua pihak yang terkait dalam penyelesaian tugas akhir termasuk Keluarga, Sahabat dan Teman-teman.

Penulis berharap dengan adanya penulisan laporan tugas akhir ini dapat memberikan manfaat ataupun referensi bagi siapapun yang membutuhkan dan bisa dijadikan sebagai pengetahuan yang berguna dalam analisis sentimen.

Yogyakarta, 19 Desember 2024

Penulis

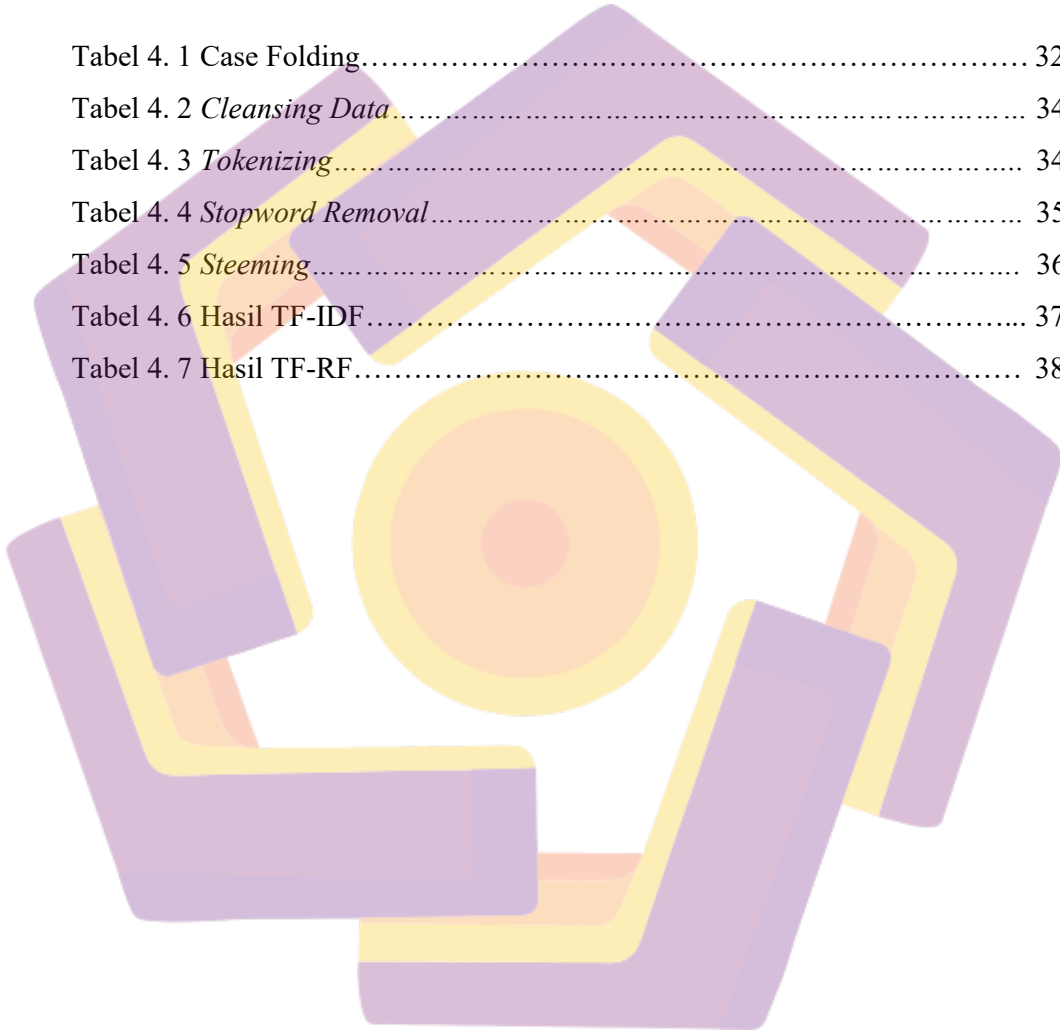
DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	17
1.1 Latar Belakang.....	17
1.2 Rumusan Masalah.....	18
1.3 Batasan Masalah	18
1.4 Tujuan Penelitian	19
1.5 Manfaat Penelitian	19
1.6 Sistematika Penulisan	19
BAB II TINJAUAN PUSTAKA	21
2.1 Studi Literatur	21
2.2 Dasar Teori	28
2.2.13 SMOTE	36
BAB III METODE PENELITIAN	40
3.1 Objek Penelitian.....	40
3.2 Alur Penelitian	40

3.2.1	Pengumpulan Data	41
3.2.2	Pre-processing Data	42
3.2.3	Labelling Dataset.....	44
3.2.4	Pembagian data uji dan data latih (splitting data)	44
3.2.5	Pembobotan kata	44
3.2.6	Balancing data (smote).....	44
3.2.7	Metode Klasifikasi	45
3.2.7.1	Multinomial Naïve Bayes Classifier.....	45
3.2.8	Evaluasi Model.....	45
BAB IV	HASIL DAN PEMBAHASAN	47
4.5.1	TF-IDF	55
4.5.2	TF-RF.....	56
4.7	Evaluasi Model	59
4.7.1	Confusion matrix	59
4.7.2	K-fold Cross Validation.....	60
BAB V	PENUTUP	62
5.1	Kesimpulan	62
5.2	Saran	62
REFERENSI		63
LAMPIRAN		68

DAFTAR TABEL

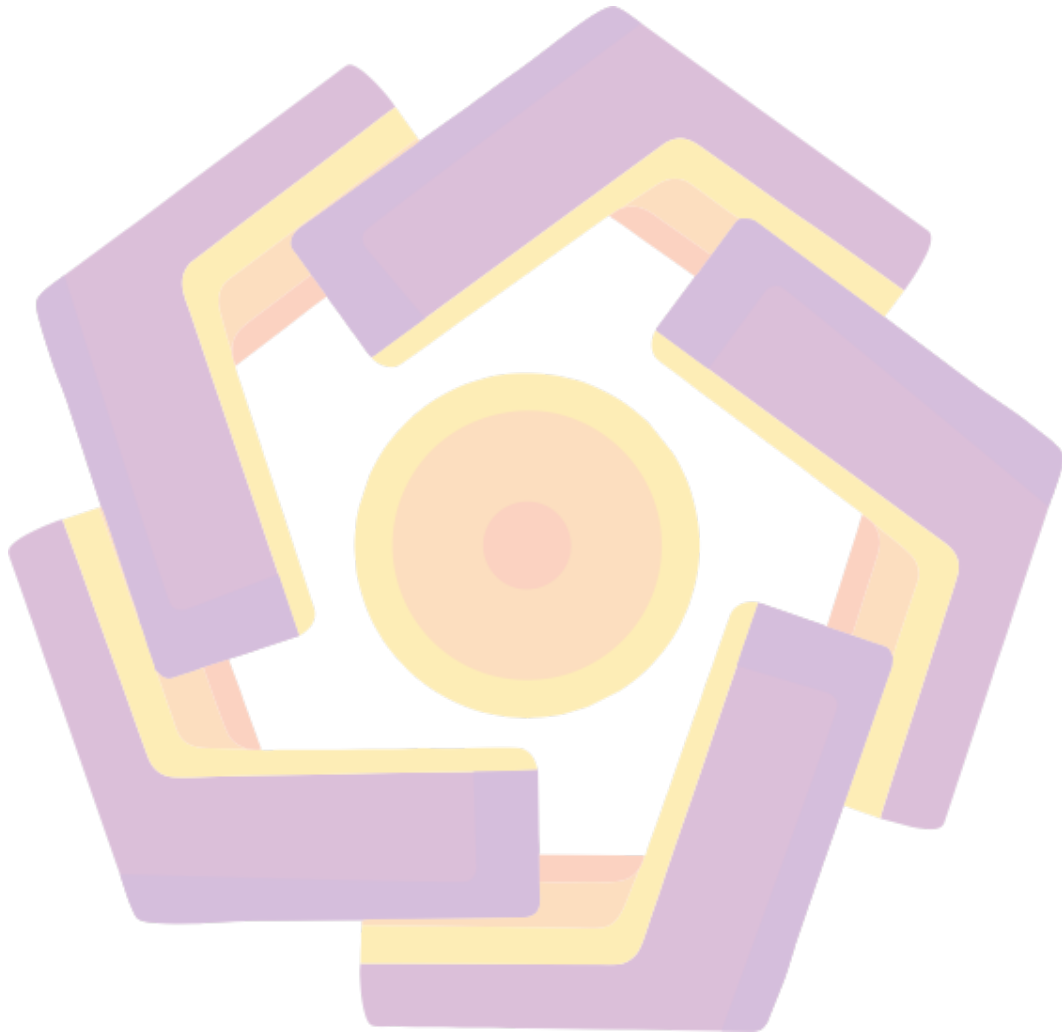
Tabel 2. 1 Keaslian Penelitian.....	8
Tabel 2. 2 Tabel <i>Confusion Matrix</i>	22
Tabel 4. 1 Case Folding.....	32
Tabel 4. 2 <i>Cleansing Data</i>	34
Tabel 4. 3 <i>Tokenizing</i>	34
Tabel 4. 4 <i>Stopword Removal</i>	35
Tabel 4. 5 <i>Steeming</i>	36
Tabel 4. 6 Hasil TF-IDF.....	37
Tabel 4. 7 Hasil TF-RF.....	38



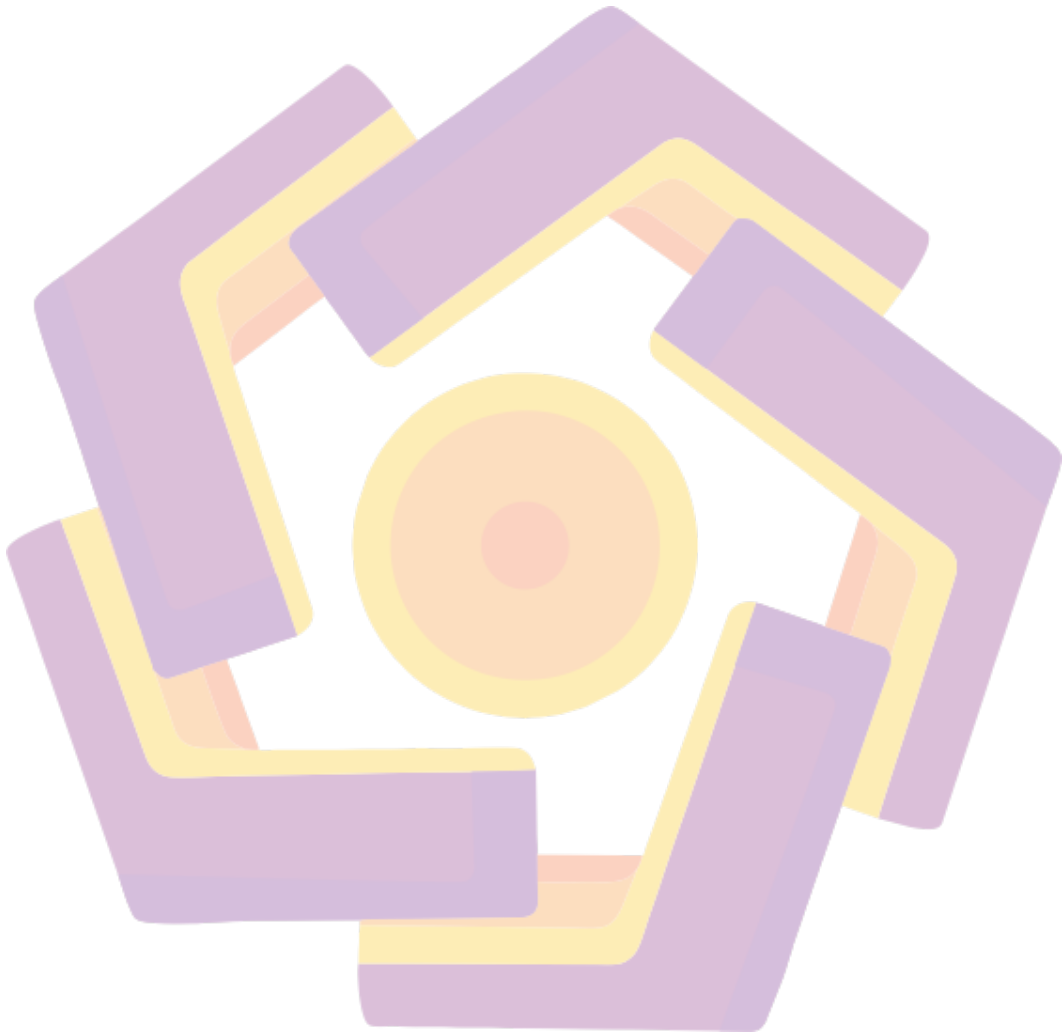
DAFTAR GAMBAR

Gambar 3. 1 <i>Flowchart</i> Penelitian	41
Gambar 3. 2 <i>Flowchart</i> Pre-processing	42
Gambar 4. 1 <i>Crawling data</i>	47
Gambar 4. 2 <i>Hasil Crawling Data</i>	48
Gambar 4. 3 <i>Hasil Crawling Data</i>	48
Gambar 4. 4 <i>Hasil Crawling Data</i>	48
Gambar 4. 5 <i>Hasil Crawling Data</i>	49
Gambar 4. 6 <i>Hasil Crawling Data</i>	49
Gambar 4. 7 <i>Case Folding</i>	50
Gambar 4. 8 <i>Coding Cleansing data</i>	51
Gambar 4. 9 <i>Cleansing data</i>	51
Gambar 4. 10 <i>Duplicate removal</i>	52
Gambar 4. 11 <i>Tokenizing</i>	52
Gambar 4. 12 <i>Stopword Removal</i>	53
Gambar 4. 13 <i>Steeming</i>	53
Gambar 4. 14 Jumlah Labelling Data per kategori	54
Gambar 4. 15 Data Splitting	54
Gambar 4. 16 Jumlah data latih	55
Gambar 4. 17 Jumlah data uji	55
Gambar 4. 18 Jumlah pelabelan data sebelum oversampling	55
Gambar 4. 19 Jumlah pelabelan data setelah oversampling	55
Gambar 4. 20 TF-IDF	56
Gambar 4. 21 TF-RF	57
Gambar 4. 22 <i>Naïve Bayes</i>	58
Gambar 4. 23 Hasil <i>Naïve Bayes</i> dengan TF-IDF	58
Gambar 4. 24 Hasil <i>Naïve Bayes</i> dengan TF-RF	58
Gambar 4. 25 Coding Confusion Matrix	59
Gambar 4. 26 Hasil Confusion Matrix TF-IDF	59
Gambar 4. 27 Hasil Confusion Matrix TF-RF	60

Gambar 4. 28 Hasil K-Fold Cross-Validation TF-IDF	60
Gambar 4. 29 K-Fold Cross-Validation TF-RF	61



DAFTAR LAMPIRAN



DAFTAR LAMBANG DAN SINGKATAN

TFIDF (Term Frequency Inverse Document Frequency)

NB (NAÏVE BAYES)

X Data dengan class yang belum diketahui

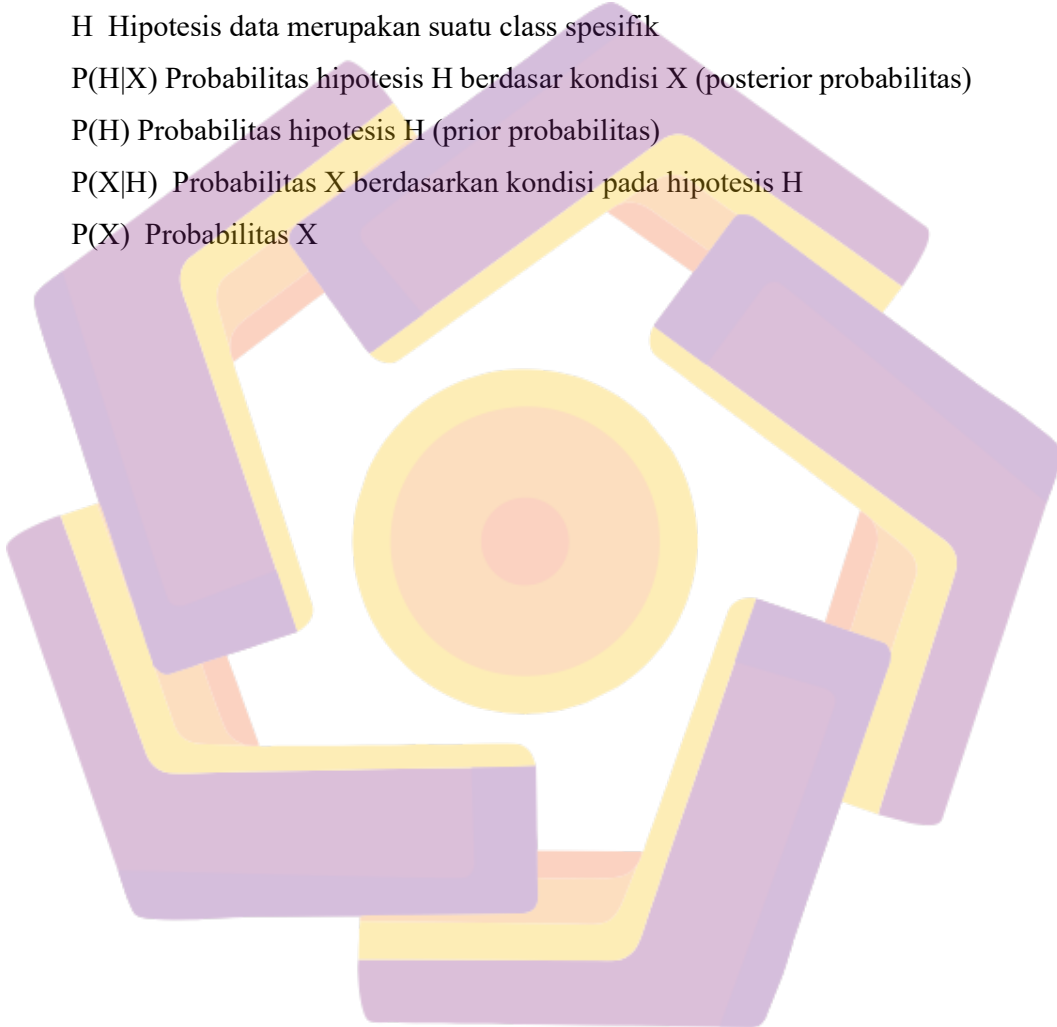
H Hipotesis data merupakan suatu class spesifik

$P(H|X)$ Probabilitas hipotesis H berdasar kondisi X (posterior probabilitas)

$P(H)$ Probabilitas hipotesis H (prior probabilitas)

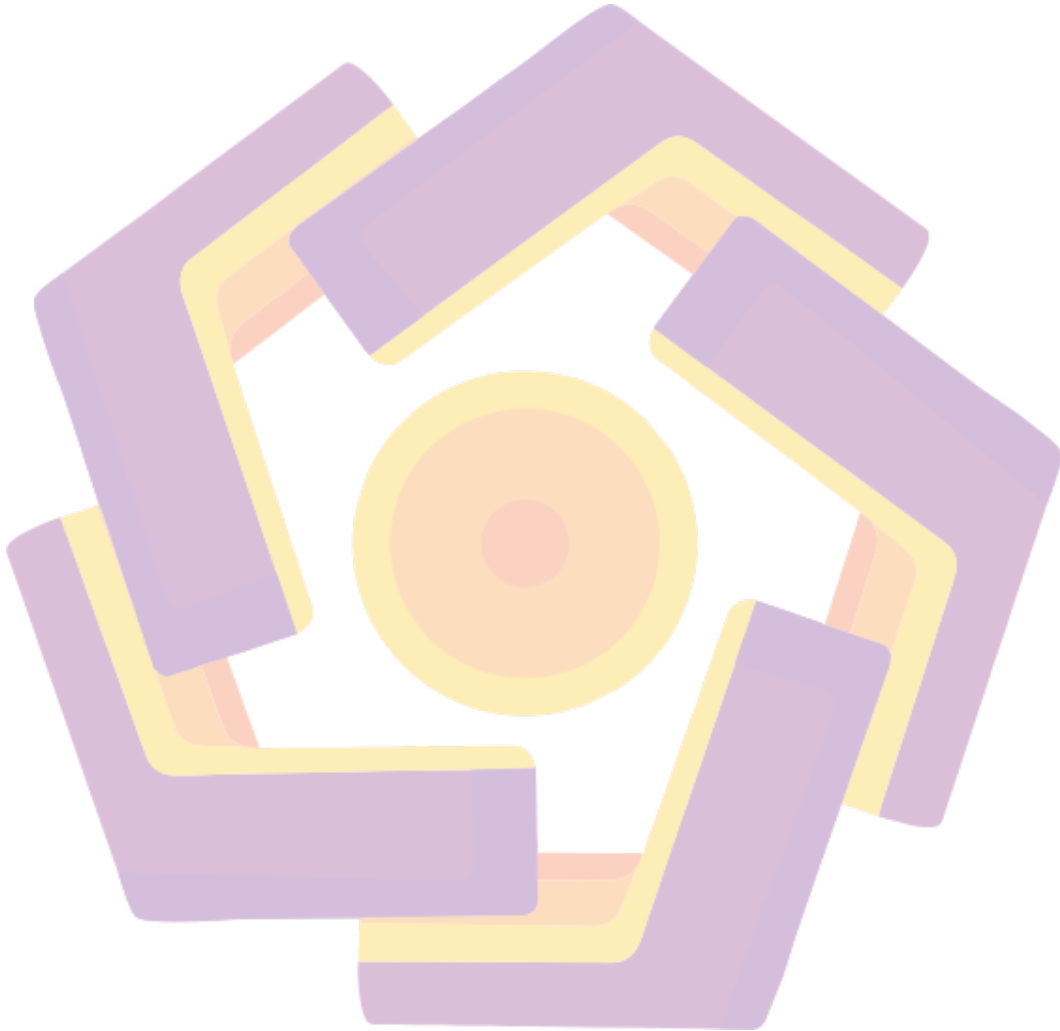
$P(X|H)$ Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$ Probabilitas X



DAFTAR ISTILAH

Vektor	besaran yang mempunyai arah
Eigen Value	akar akar persamaan



INTISARI

Kemajuan dunia digital memudahkan akses informasi melalui portal dan media sosial. Media sosial menjadi platform utama bagi opini publik, mencakup pujian, ujaran kebencian, dan hoaks, memicu perdebatan online. X menjadi medium diskusi yang umum digunakan.

Penerapan analisis sentimen di X memungkinkan pengguna mendapatkan kritik dan rekomendasi yang dapat diandalkan, seperti pengesahan RUU KUHP oleh DPR RI. Pendapat yang terkumpul memberikan kontribusi positif terhadap keputusan pihak terkait.

Penelitian ini bertujuan menganalisis sejauh mana opini masyarakat terkait RUU KUHP yang baru disahkan. Analisis sentimen menggunakan Metode Klasifikasi Naive Bayes (NBC), teknik pembelajaran mesin berbasis probabilitas yang terkenal akan akurasi dan efisiensinya.

RUU KUHP mencerminkan keseimbangan antara kepentingan pribadi dan masyarakat, serta antara kepastian hukum dan keadilan. Pengesahan RUU tersebut pada 12 Juni 2022 oleh DPR RI menciptakan polarisasi di masyarakat dengan beragam perspektif. Analisis sentimen di X bertujuan memahami opini publik terhadap isu ini.

Dengan menggunakan NBC, penelitian ini mengklasifikasikan tweet ke dalam sentimen negatif, positif, atau netral, yang diambil secara online melalui tweet harvest. Hasil penelitian menggunakan metode NBC dengan pembobotan TF-IDF menghasilkan akurasi adalah 80%. Sedangkan, Hasil penelitian menggunakan metode NBC dengan pembobotan TF-RF menghasilkan akurasi adalah 83%.

Kata kunci: Analisis Sentimen, Naive Bayes, Tweet Harvest.

ABSTRACT

The advancing digital world facilitates easy access to information through portals and social media. Social media serves as a prominent platform for public opinions, encompassing praise, hate speech, and hoaxes, sparking online debates. X is a commonly used discussion medium.

Utilizing sentiment analysis on X enables users to acquire reliable critiques and recommendations, such as the approval of the Criminal Code Bill by the Indonesian Parliament. Cumulative opinions contribute positively to decisions made by relevant authorities.

The study aims to analyze the extent of public opinions on the recently enacted RUU KUHP. Sentiment analysis employs the Naive Bayes Classifier (NBC), a probabilistic machine learning technique known for its accuracy and efficiency of RUU KUHP reflects a balance between personal and societal interests and legal certainty and justice. Its approval on June 12, 2022, by the Indonesian Parliament has led to polarization in society, with varied perspectives. X sentiment analysis seeks to comprehend public opinions on this issue.

Using NBC, this research classifies tweets into negative, positive, or neutral sentiments, which are taken online through tweet harvest. The results of research using the NBC method with TF-IDF treatment resulted in an accuracy of 80%. Meanwhile, the results of research using the NBC method with TF-RF treatment resulted in an accuracy of 83%.

Keyword: Sentiment Analysis, Naive Bayes, Tweet Harvest.