

**PENDEKATAN MULTINOMIAL *NAÏVE BAYES* UNTUK
KLASIFIKASI JENIS-JENIS *CYBER HARASSMENT* DI
MEDIA SOSIAL TWITTER**

SKRIPSI

**OLEH:
KIKI OKTAVIA
NIM. 19610079**



**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2023**

**PENDEKATAN MULTINOMIAL *NAÏVE BAYES* UNTUK
KLASIFIKASI JENIS-JENIS *CYBER HARASSMENT* DI
MEDIA SOSIAL TWITTER**

SKRIPSI

**Diajukan Kepada
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
untuk Memenuhi Salah Satu Persyaratan
dalam Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh
Kiki Oktavia
NIM. 19610079**

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2023**

**PENDEKATAN MULTINOMIAL *NAÏVE BAYES* UNTUK
KLASIFIKASI JENIS-JENIS *CYBER HARASSMENT* DI
MEDIA SOSIAL TWITTER**

SKRIPSI

**Oleh
Kiki Oktavia
NIM. 19610079**

Telah Disetujui Untuk Diuji
Malang, 26 September 2023

Dosen Pembimbing I



Ria Dhea Layla N. K., M.Si.
NIP. 19900709 202321 2 037


Dosen Pembimbing II



Ari Kusumastuti, M.Pd., M.Si.
NIP. 19770521 200501 2 004

Mengetahui,
Ketua Program Studi Matematika




Dr. Elly Susanti, M.Sc.
NIP. 19741129 200012 2 005

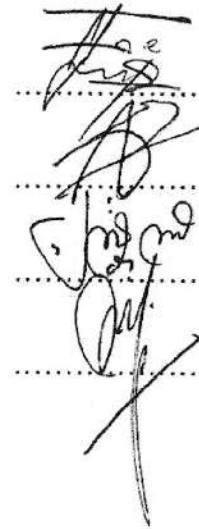
**PENDEKATAN MULTINOMIAL *NAÏVE BAYES* UNTUK
KLASIFIKASI JENIS-JENIS *CYBER HARASSMENT* DI
MEDIA SOSIAL TWITTER**

SKRIPSI

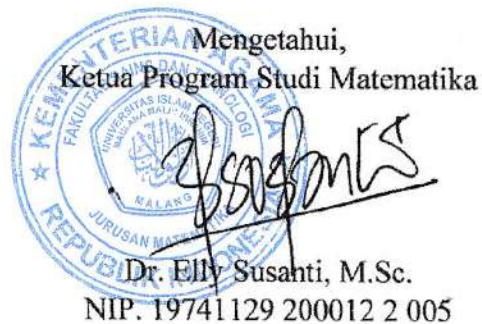
**Oleh
Kiki Oktavia
NIM. 19610079**

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)
Tanggal, 02 Oktober 2023

Ketua Penguji : Fachrur Rozi, M.Si.
Anggota Penguji 1 : Angga Dwi Mulyanto, M.Si.
Anggota Penguji 2 : Ria Dhea Layla N. K., M.Si.
Anggota Penguji 3 : Ari Kusumastuti, M.Pd., M.Si.



Mengetahui,
Ketua Program Studi Matematika



Dr. Elly Susanti, M.Sc.
NIP. 19741129 200012 2 005

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Kiki Oktavia

NIM : 19610079

Program Studi : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Pendekatan Multinomial *Naïve Bayes* untuk Klasifikasi Jenis-Jenis *Cyber Harassment* di Media Sosial Twitter

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan dan pikiran saya, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka. Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perilaku tersebut.

Malang, 02 Oktober 2023
Yang membuat pernyataan,



Kiki Oktavia

NIM. 19610079

MOTO

“ ... Dan janganlah kamu membunuh dirimu; sungguh, Allah Maha Penyayang kepadamu.”

(Q.S. An-Nisa:29)

"Que Será, Será (Whatever Will Be, Will Be)"

(Doris Day, 1956)

PERSEMBAHAN

Skripsi ini penulis persembahkan kepada:
Allah *subhanahu wa ta'ala* Yang Maha Pengasih,
almarhum Bapak yang selalu penulis rindukan,
Mama yang selalu merestui langkah penulis,
kakak-kakak penulis yang selalu bekerja keras,
dan kepada diri penulis sendiri.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Segala puji bagi Allah *subhanahu wa ta'ala* yang telah melimpahkan rahmat, taufik, serta hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul “Pendekatan Multinomial *Naïve Bayes* untuk Klasifikasi Jenis-Jenis *Cyber Harassment* di Media Sosial Twitter”. Judul tersebut diangkat oleh penulis berdasarkan ketertarikan penulis terhadap topik dan penerapan ilmu matematika pada kehidupan nyata. Penelitian ini mengkaji mengenai klasifikasi jenis-jenis *Cyber Harassment* di media sosial Twitter dengan metode Multinomial *Naïve Bayes*. Selesaiannya skripsi ini tidak lepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Prof. Dr. H. M. Zainuddin, M.A., selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Prof. Dr. Hj. Sri Harini, M.Si., selaku dekan Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Elly Susanti, M.Sc., selaku ketua Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang
4. Ria Dhea Layla Nur Karisma, M.Si., selaku dosen pembimbing 1 yang membimbing penulis dengan memberikan pengetahuan, arahan, nasihat, dukungan, hingga pengertian agar penulis dapat menyelesaikan skripsi dengan baik.
5. Ari Kusumastuti, M.Pd., M.Si., selaku dosen pembimbing 2 yang membimbing penulis dengan memberikan pengetahuan dan saran agar penulis dapat menyelesaikan skripsi dengan baik.
6. Fachrur Rozi, M.Si., selaku ketua penguji dalam sidang skripsi yang telah memberikan ilmu pengetahuan dan saran yang bermanfaat bagi penulis
7. Angga Dwi Mulyanto, M.Si., selaku anggota penguji 1 dalam sidang skripsi yang telah memberikan ilmu pengetahuan dan saran yang bermanfaat bagi penulis.

8. Seluruh dosen dan staf Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim yang telah memberikan ilmu dan bantuan selama perkuliahan.
9. Almarhum Bapak yang selalu berada dalam kenangan, Ibu yang selalu lebih semangat dari siapa pun dalam mendukung penulis, Kakak perempuan penulis yang telah bekerja keras untuk menghidupi penulis serta keluarga, Kakak-kakak lelaki penulis yang juga mendukung penulis, dan seluruh keluarga yang telah membantu baik secara materil maupun spiritual.
10. Teman-teman terdekat penulis, Naftalie Mahardika, Camelia Sugeng Saputri, Arnetta Dinda Aulia Puspita, Maulidyah Juanita, Alwiyah Muharromi, dan Nurul Fatimah.
11. Seluruh karya sastra yang telah menemani dan mengisi waktu luang penulis.
12. Mahasiswa Program Studi Mahasiswa, terkhusus teman-teman angkatan 2019, yang selalu saling menolong, menyemangati, dan memudahkan satu sama lain.
13. Semua pihak yang tidak dapat disebutkan satu per satu yang turut mendoakan, mendukung, dan meyakinkan penulis untuk segera menyelesaikan skripsi.

Semoga Allah *subhanahu wa ta'ala* memberikan balasan atas kebaikan dan ketulusan yang telah diberikan kepada penulis. Penulis menyadari masih banyaknya kekurangan dalam penyusunan skripsi ini, sehingga penulis menerima kritik dan saran yang bersifat membangun dari semua pihak. Penulis berharap, skripsi ini dapat bermanfaat setidaknya bagi penulis, dan bagi para pembaca, serta dapat menjadi referensi dan menambah ilmu pengetahuan.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Malang, 02 Oktober 2023

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGANTAR	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTO.....	vi
PERSEMBAHAN	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR SIMBOL	xiv
DAFTAR LAMPIRAN.....	xv
ABSTRAK	xvi
ABSTRACT.....	xvii
مستخلص البحث	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Tujuan Penelitian	7
1.4 Manfaat Penelitian	7
1.5 Batasan Masalah.....	7
BAB II TINJAUAN PUSTAKA	9
2.1 Teorema Bayes	9
2.1.1 Naïve Bayes Classifier	10
2.1.2 Multinomial Naïve Bayes	12
2.2 Twitter.....	19
2.3 Cyber Harassment.....	20
2.3.1 Physical Threats	21
2.3.2 Purposeful Embarrassment	22
2.3.3 Racist.....	22
2.3.4 Sexual Harassment.....	23
2.4 Data Crawling	23
2.5 Data Preprocessing	24
2.6 Ekstraksi Fitur	26
2.7 Uji Performa Model	26
2.8 Cyber Harassment dalam Pandangan Islam.....	27
2.9 Kajian Cyber Harassment dengan Teori Pendukung.....	29
BAB III METODE PENELITIAN	33
3.1 Pendekatan Penelitian	33
3.2 Teknik Pengumpulan Data	33
3.3 Jenis dan Sumber Data.....	34
3.4 Variabel Penelitian	35
3.5 Tahapan Penelitian.....	36
3.6 Diagram Alir.....	38
BAB IV HASIL DAN PEMBAHASAN	39

4.1	<i>Data Crawling</i>	39
4.2	<i>Data Preprocessing</i>	41
4.3	Analisis Deskriptif	42
4.3.1	Analisis Deskriptif Variabel Independen (x)	42
4.3.2	Analisis Deskriptif Variabel Dependen (y)	45
4.4	Ekstraksi Fitur	52
4.5	Analisis Metode Multinomial <i>Naïve Bayes</i>	53
4.5.1	Pembagian Data <i>Training</i> dan Data <i>Testing</i>	53
4.5.2	Klasifikasi Jenis-Jenis <i>Cyber Harassment</i> Menggunakan Metode Multinomial <i>Naïve Bayes</i>	54
4.5.3	Hasil Klasifikasi Jenis-Jenis <i>Cyber Harassment</i> Menggunakan Metode Multinomial <i>Naïve Bayes</i>	61
4.5.4	Ketepatan Klasifikasi	62
4.6	Uji Performa Model	65
4.7	Kajian Keislaman dengan Hasil Penelitian	66
BAB V PENUTUP		69
5.1	Kesimpulan	69
5.2	Saran	70
DAFTAR PUSTAKA		71
LAMPIRAN		76
RIWAYAT HIDUP		94

DAFTAR TABEL

Tabel 2.1	Tabulasi Silang Klasifikasi <i>Multiclass</i>	18
Tabel 2.2	Interval Kategori Nilai Akurasi	19
Tabel 2.3	Contoh Kalimat Dalam Dokumen.....	26
Tabel 2.4	Frekuensi Kemunculan Kata Dalam Dokumen	26
Tabel 3.1	Kata Kunci	34
Tabel 3.2	Data Penelitian.....	34
Tabel 3.3	<i>Tweet</i> dengan Pelabelan Kelas.....	35
Tabel 3.4	Variabel <i>Cyber Harassment</i>	36
Tabel 4.1	Hasil <i>Data Crawling</i> dan Pemberian Kelas	40
Tabel 4.2	Hasil <i>Data Preprocessing</i>	42
Tabel 4.3	Variabel Independen (x)	43
Tabel 4.4	Hasil Ekstraksi Fitur dengan <i>Bag of Words</i>	52
Tabel 4.5	Pembagian <i>Data Training</i> dan <i>Data Testing</i>	53
Tabel 4.6	Contoh <i>Dataset</i> untuk Klasifikasi	54
Tabel 4.7	Probabilitas <i>Prior</i> Contoh <i>Dataset</i> untuk Klasifikasi	55
Tabel 4.8	Jumlah Kata Pada Setiap Kelas	56
Tabel 4.9	Jumlah Kata Berbeda (<i>Vocabulary</i>)	56
Tabel 4.10	Ekstraksi Fitur Dokumen 5	57
Tabel 4.11	Perhitungan Nilai <i>Likelihood</i> $P(W_i C)$	58
Tabel 4.12	Probabilitas Bersyarat	60
Tabel 4.13	Hasil Klasifikasi <i>Data Testing</i> Menggunakan <i>Multinomial Naïve Bayes</i>	61
Tabel 4.14	Jumlah Hasil Klasifikasi <i>Cyber Harassment</i> Menggunakan <i>Multinomial Naïve Bayes</i> Pada Setiap Kelas	62
Tabel 4.15	Hasil Tabulasi Silang	62
Tabel 4.16	Hasil <i>K-fold Cross Validation</i>	65

DAFTAR GAMBAR

Gambar 2.1	<i>Bag of Words Assumption</i> pada Multinomial Naïve Bayes	13
Gambar 2.2	Alur <i>Data Preprocessing</i>	25
Gambar 3.1	Diagram Alir Penelitian	38
Gambar 4.1	Frekuensi Kemunculan Kata Tertinggi Pada Kasus <i>Cyber Harassment</i>	43
Gambar 4.2	<i>Wordcloud</i> Pada Kasus <i>Cyber Harassment</i>	44
Gambar 4.3	Frekuensi Kemunculan Kata Tertinggi pada Kelas <i>Physical Threats</i>	45
Gambar 4.4	<i>Wordcloud</i> Pada Kelas <i>Physical Threats</i>	46
Gambar 4.5	Frekuensi Kemunculan Kata Tertinggi pada Kelas <i>Purposeful Embarrassment</i>	47
Gambar 4.6	<i>Wordcloud</i> Pada Kelas <i>Purposeful Embarrassment</i>	48
Gambar 4.7	Frekuensi Kemunculan Kata Tertinggi pada Kelas <i>Racist</i>	49
Gambar 4.8	<i>Wordcloud</i> Pada Kelas <i>Racist</i>	50
Gambar 4.9	Frekuensi Kemunculan Kata Tertinggi pada Kelas <i>Sexual Harassment</i>	50
Gambar 4.10	<i>Wordcloud</i> Pada Kelas <i>Sexual Harassment</i>	52

DAFTAR SIMBOL

$P(C D)$: Probabilitas <i>posterior</i> kelas C benar terjadi dengan diberikan dokumen D
$P(C)$: Probabilitas <i>prior</i> dari kelas C
$P(W_i C)$: Probabilitas kata ke- i pada dokumen dengan diberikan kelas C (<i>likelihood</i>)
D	: Dokumen
C	: Kelas
W_i	: Kata ke- i
W_1, W_2, \dots, W_n	: Barisan kata ke-1, 2, ..., n
c_i	: Sub-kelas c ke- i
n_{c_i}	: Jumlah kelas c ke- i
N_d	: Total keseluruhan kelas
x_{c_i}	: Jumlah kemunculan kata ke- i pada kelas c
$x_{c_1}, x_{c_2}, \dots, x_{c_n}$: Jumlah kemunculan kata ke-1, 2, ..., n pada kelas c
N	: Jumlah seluruh kata yang muncul
θ_{c_i}	: Probabilitas kemunculan kata ke- i pada kelas c
L	: Fungsi <i>likelihood</i>
\ln	: Logaritma natural
ℓ	: <i>Log-likelihood</i>
$\frac{\partial \ell}{\partial \theta_c}$: Persamaan diferensial dari $\ln L(\theta_{c_i})$
$\widehat{\theta}_{c_i}$: <i>Maximum Likelihood Estimation</i> θ_{c_i}
α	: <i>Add-one smoothing</i> untuk setiap kata pada kelas c
α_d	: Total seluruh kata berbeda pada seluruh dokumen (<i>vocabulary</i>)
T_{mn}	: <i>True class</i> pada baris ke- m dan kolom ke- n pada tabulasi silang
F_{mn}	: <i>False class</i> pada baris ke- m dan kolom ke- n pada tabulasi silang
k	: Jumlah <i>fold</i>

DAFTAR LAMPIRAN

Lampiran 1. Data <i>Tweet</i> Kasus <i>Cyber Harassment</i>	76
Lampiran 2. Hasil <i>Data Preprocessing</i>	78
Lampiran 3. Hasil Ekstraksi Fitur	82
Lampiran 4. Hasil Klasifikasi Jenis-Jenis <i>Cyber Harassment</i> Menggunakan Multinomial <i>Naïve Bayes</i>	84
Lampiran 5. <i>Script Data Crawling</i>	89
Lampiran 6. <i>Script Data Preprocessing</i>	90
Lampiran 7. <i>Script Multinomial Naïve Bayes</i>	93

ABSTRAK

Oktavia, Kiki. 2023. **Pendekatan Multinomial Naïve Bayes untuk Klasifikasi Jenis-Jenis Cyber Harassment di Media Sosial Twitter**. Skripsi. Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: I) Ria Dhea Layla Nur Karisma, M.Si., II) Ari Kusumastuti, M.Pd., M.Si.

Kata Kunci: Multinomial Naïve Bayes, Teorema Bayes, Cyber Harassment, Jenis-Jenis Cyber Harassment, Twitter

Multinomial Naïve Bayes merupakan metode klasifikasi dari Naïve Bayes Classifier yang berlandaskan pada Teorema Bayes dan distribusi multinomial yang bekerja dengan optimal pada kasus klasifikasi *multiclass* data teks. Multinomial Naïve Bayes menghitung peluang kemunculan setiap kata dengan mengalikan nilai probabilitas *prior* kelas dengan nilai *likelihood* kemunculan setiap kata pada setiap kelas. Fenomena *Cyber Harassment* didefinisikan sebagai perilaku memanfaatkan teknologi dengan tujuan untuk merugikan atau mempermalukan seseorang. Twitter merupakan salah satu media sosial dengan kasus *Cyber Harassment* tertinggi. Data yang digunakan dalam penelitian ini adalah 2.000 *tweet* berbahasa Indonesia yang dikumpulkan menggunakan *Application Programming Interface* yang disediakan oleh Twitter pada bulan Januari hingga Juni 2023. Jenis-jenis *Cyber Harassment* yang terdiri dari *Physical Threats*, *Purposeful Embarrassment*, *Racist*, dan *Sexual Harassment* merupakan variabel dependen, sedangkan kata-kata yang terkumpul pada *tweet* adalah variabel independen. Penelitian ini menggunakan 95% data *training* dan 5% data *testing*. Penelitian ini bertujuan untuk mendapatkan hasil klasifikasi dengan akurasi yang baik mengenai jenis-jenis *Cyber Harassment* di media sosial Twitter menggunakan metode Multinomial Naïve Bayes. Hasil klasifikasi yang diperoleh yaitu 20 dokumen terklasifikasi sebagai *Physical Threats*, 10 dokumen terklasifikasi sebagai *Purposeful Embarrassment*, 25 dokumen terklasifikasi sebagai *Racist*, dan 22 dokumen terklasifikasi sebagai *Sexual Harassment*. Tingkat akurasi klasifikasi jenis-jenis *Cyber Harassment* di media sosial Twitter menggunakan Multinomial Naïve Bayes sebesar 77% dan hasil uji performa model dengan *K-fold Cross Validation* sebesar 76,21% menunjukkan bahwa metode Multinomial Naïve Bayes dapat mengklasifikasi jenis-jenis *Cyber Harassment* di media sosial Twitter dengan baik.

ABSTRACT

Oktavia, Kiki. 2023. **Multinomial Naïve Bayes Approach to Classify Types of Cyber Harassment on Social Media Twitter**. Thesis. Department of Mathematics, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Advisors: I) Ria Dhea Layla Nur Karisma, M.Si., II) Ari Kusumastuti, M.Pd., M.Si.

Keywords: Multinomial Naïve Bayes, Bayes Theorem, Cyber Harassment, Types of Cyber Harassment, Twitter

Multinomial Naïve Bayes is a classification method in Naïve Bayes Classifier based on Bayes Theorem and multinomial distribution that works optimally in the multiclass classification of text data. Furthermore, it calculates the probability of occurrence of each word by multiplying the class prior probability by the likelihood value of the occurrence of each word in each class. The phenomenon of Cyber Harassment is defined as the behavior of utilizing technology to harm or humiliate people. Twitter is one of the social media with the highest Cyber Harassment cases. The data used in this study is collected 2,000 Indonesian tweets using Application Programming Interface provided by Twitter from January to June 2023. The types of Cyber Harassment, which consist of Physical Threats, Purposeful Embarrassment, Racist, and Sexual Harassment, are the dependent variables, and the words collected are the independent variables. Then, it used 95% training data and 5% testing data. This study aims to classify results accurately regarding the types of Cyber Harassment on Twitter using the Multinomial Naïve Bayes method. The classification results obtained are 20 documents classified as Physical Threats, 10 documents classified as Purposeful Embarrassment, 25 documents classified as Racist, and 22 documents classified as Sexual Harassment. The accuracy of classification of types of Cyber Harassment on Twitter using Multinomial Naïve Bayes is 77% and the results of the model performance test with K-fold Cross Validation is 76.21%, showing that the Multinomial Naïve Bayes method can classify the types of Cyber Harassment on Twitter effectively.

مستخلص البحث

أوكتافيا، كيكبي ٢٠٢٣. نُحج بايز الساذج متعدد الجنسيات لتصنيف أنواع التحرش الإلكتروني على وسائل التواصل الاجتماعي على تويتر. أطروحة. قسم الرياضيات، كلية العلوم والتكنولوجيا، مولانا مالك إبراهيم الدولة الإسلامية جامعة مالانج. المستشارون: (١) ،راي داي ليلي نور كر مسالما، الماجستير.، (٢) اري كوسوماسطوتي، الماجستير.

الكلمات الرئيسية: ساذجة بايز متعددة الجنسيات، نظرية بايز، التحرش الإلكتروني، أنواع المضايقات الإلكترونية ، تويتر

Naïve Bayes متعدد الحدود (MNB) هي طريقة تصنيف في Naïve Bayes Classifier تعتمد على نظرية بايز والتوزيع متعدد الحدود الذي يعمل على النحو الأمثل في التصنيف متعدد الفئات للبيانات النصية. علاوة على ذلك ، فإنه يحسب احتمال حدوث كل كلمة بضرب الاحتمال السابق للفئة في قيمة احتمال حدوث كل كلمة في كل فئة. تعرف ظاهرة التحرش السيرياني بأنها سلوك استخدام التكنولوجيا لإيذاء الناس أو إذلالهم. تويتر هو واحد من وسائل التواصل الاجتماعي مع أعلى حالات التحرش السيرياني. تم جمع البيانات المستخدمة في هذه الدراسة 2000 تغريدة إندونيسية باستخدام واجهة برمجة التطبيقات التي يوفرها Twitter من يناير إلى يونيو 2023. أنواع التحرش السيرياني ، والتي تتكون من التهديدات الجسدية ، والإحراج المتعمد ، والعنصرية ، والتحرش الجنسي ، هي المتغيرات التابعة ، والكلمات التي تم جمعها هي المتغيرات المستقلة. بعد ذلك ، استخدمت 95٪ من بيانات التدريب و 5٪ من بيانات الاختبار. تهدف هذه الدراسة إلى تصنيف النتائج بدقة فيما يتعلق بأنواع التحرش الإلكتروني على تويتر باستخدام طريقة MNB. نتائج التصنيف التي تم الحصول عليها هي 20 وثيقة مصنفة على أنها تهديدات جسدية ، و 10 وثائق مصنفة على أنها إحراج متعمد ، و 25 وثيقة مصنفة على أنها عنصرية ، و 22 وثيقة مصنفة على أنها تحرش جنسي. تبلغ دقة تصنيف أنواع التحرش الإلكتروني على تويتر باستخدام ساذج بايز متعدد الحدود 77٪. ونتائج اختبار أداء النموذج باستخدام التحقق من صحة K-fold Cross هي 76.21٪ ، مما يدل على أن طريقة MNB يمكنها تصنيف أنواع التحرش الإلكتروني على تويتر بشكل فعال.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan serta kemajuan yang pesat telah terjadi di berbagai rumpun ilmu, di mana perkembangan tersebut membawa berbagai dampak dan perubahan dalam kehidupan manusia. Dalam bidang Sains dan Teknologi, statistika memiliki peran penting dalam menguji hingga menganalisis derajat akurasi dari suatu hipotesis dalam sebuah eksperimen. Melalui berbagai distribusi peluang dalam ilmu statistika yang kemudian dapat diimplementasikan untuk menyelesaikan permasalahan pada kasus klasifikasi hingga model *multiclass*. Klasifikasi *multiclass* merupakan model klasifikasi yang mempunyai setidaknya dua hasil atau *ouput* yang berbeda, dengan tujuan untuk mendapatkan kelas yang tepat untuk data yang diberikan.

Permasalahan pada kasus klasifikasi *multiclass* dapat diatasi dengan berbagai jenis metode, di antaranya yang memiliki kinerja model yang baik adalah Multinomial *Naïve Bayes*. Multinomial *Naïve Bayes* (MNB) adalah metode yang juga bekerja dengan optimal pada kasus klasifikasi *multiclass* khususnya pada data teks, di mana data penelitian menggunakan jumlah frekuensi kata dari hasil ekstraksi fitur. MNB memiliki algoritma yang sederhana tetapi dapat menghasilkan tingkat akurasi klasifikasi yang baik dan meminimalisir nilai eror. Namun, hasil akurasi yang dimiliki oleh metode ini memiliki faktor penghambat dikarenakan adanya “frekuensi nol” dalam perhitungan probabilitasnya (Hosein dan Baboolal, 2022).

Multinomial *Naïve Bayes* merupakan satu dari tiga jenis metode klasifikasi *Naïve Bayes Classifier* yang berlandaskan pada Teorema *Bayes* sebagai dasar perhitungannya. Multinomial *Naïve Bayes* mengklasifikasi kelas yang tepat dengan mengalikan nilai probabilitas *prior* dengan nilai *likelihood* untuk setiap kelas. Multinomial *Naïve Bayes* adalah metode yang bertujuan untuk mengklasifikasi hasil yang tepat dengan kemungkinan hasil atau *output* lebih dari dua dengan menghitung nilai probabilitas berdasarkan pada distribusi multinomial (Reiten, 2017). Pada era kemajuan Sains dan Teknologi saat ini, Multinomial *Naïve Bayes* dapat diimplementasikan dalam bidang *machine learning* atau *text mining* untuk menemukan informasi baru. Dalam kasus klasifikasi data teks, Multinomial *Naïve Bayes* bekerja lebih baik dari Gaussian maupun Bernoulli *Naïve Bayes*, dikarenakan metode ini menghitung probabilitas dari frekuensi untuk setiap kata yang terdapat pada dokumen.

Faktanya, kemajuan pada bidang Sains dan Teknologi juga memberi dampak negatif dalam kehidupan manusia, salah satunya adalah kemunculan media sosial. Kebebasan bermedia sosial yang tidak dapat dikendalikan menjadi celah yang dimanfaatkan oleh sejumlah oknum untuk melakukan perilaku menyimpang atau kejahatan di ruang siber (*cybercrime*). Salah satu jenis dari *Cybercrime* adalah *Cyber Harassment* (pelecehan *online*), yang didefinisikan sebagai perilaku yang disengaja dan berulang dengan memanfaatkan teknologi seperti ponsel, *e-mail*, situs web, dan sejenisnya yang bertujuan untuk merugikan atau mempermalukan seseorang (Stevens, dkk., 2021). Pew Research Center dalam penelitiannya yang berjudul "*The State of Online Harassment*" (2021) memaparkan bahwa *cyber harassment* diklasifikasikan menjadi empat jenis, yaitu ancaman kekerasan fisik

(*physical threats*), mengejek/mempermalukan (*purposeful embarrassment*), panggilan yang menyinggung (*racist*), dan pelecehan seksual (*sexual harassment*).

Direktorat Tindak Pidana Siber Bareksrim Polri melalui portal Patrolisiber mengungkapkan terdapat 15.152 laporan atau pengaduan mengenai kejahatan siber di Indonesia sepanjang tahun 2021, di mana 3.101 laporan di antaranya merupakan kasus pelecehan dengan pengancaman dan pencemaran nama baik (Dihni, 2021). Southeast Asia Freedom of Expression Network (SAFE-net) turut menyatakan bahwa terjadi peningkatan kasus *cyber harassment* pada tahun 2020 sebesar sepuluh kali lipat dibandingkan tahun sebelumnya. Komisi Nasional Anti Kekerasan terhadap Perempuan (2022) memaparkan data berdasarkan survei yang dilakukan oleh Koalisi Ruang Publik Aman (KRPA) tahun 2021 bahwa *cyber harassment* paling banyak terjadi di media sosial (42%) dan aplikasi *chat* (33%). Indonesia sendiri menempati peringkat kelima di dunia sebagai negara dengan masyarakat yang paling banyak menggunakan media sosial Twitter, dengan jumlah 18,54 juta pengguna (Rizaty, 2022).

Sejalan dengan pemaparan pada poin sebelumnya, penelitian yang berjudul “*Crime and its fear in social media*” yang dilakukan oleh Prieto, dkk., (2020) menyatakan bahwa salah satu media sosial dengan tingkat kejahatan siber tertinggi adalah Twitter. Pada media sosial Twitter, setidaknya terdapat lima *tweet* atau unggahan setiap menitnya yang mengandung unsur kejahatan. Permasalahan ini menjadikan pemerintah mengambil langkah dengan pembentukan polisi virtual sebagai pengawasan aktivitas di media sosial yang dibentuk oleh Direktorat Tindak Pidana Siber Bareskrim Polri bersama Kementerian Komunikasi dan Informatika (Latiefah, dkk., 2022). Dalam perspektif Islam, di dalam Al-Qur’an telah

ditegaskan bahwa segala bentuk kejahatan atau kemunkaran adalah suatu perbuatan buruk dalam syariat. Allah *subhanahu wa ta'ala* berfirman dalam surah al-Nahl ayat 90 yang artinya:

“Sesungguhnya Allah menyuruh (kamu) berlaku adil dan berbuat kebajikan, memberi kepada kaum kerabat, dan Allah melarang dari perbuatan keji, kemungkaran dan permusuhan. Dia memberi pengajaran kepadamu agar kamu dapat mengambil pelajaran.”

Munkar memiliki makna sebagai segala sesuatu yang buruk, bertentangan dengan syariat, dilarang atau diharamkan, juga tidak disukai. *Munkar* juga dapat dimaknai sebagai sesuatu yang menggelisahkan hati (Muhammadun, 2011). Maka perilaku pelecehan yang merupakan perilaku buruk dan membawa dampak merugikan bagi para korbannya merupakan salah satu bentuk kemungkaran manusia. Sehingga perilaku pelecehan perlu dihilangkan dan jika ada manusia yang melakukannya maka harus diberi hukuman yang seadil-adilnya.

Penelitian terdahulu yang dilakukan oleh Arnisha Akhter, dkk., (2019) dengan judul *“Cyber Bullying Detection and Classification using Multinomial Naïve Bayes and Fuzzy Logic”* bertujuan untuk mengklasifikasikan perilaku *bullying* pada media sosial Facebook dalam tiga kelas, yaitu memermalukan (*shaming*), melecehkan (*harassment*), dan rasis (*racism*) menggunakan 1.000 data berbahasa Inggris dengan partisi 80% *data training* dan 20% *data testing*, menghasilkan tingkat akurasi yang baik sebesar 88,89%. Penelitian lain dilakukan oleh Chita Naully Harahap, dkk., (2021) dengan judul *“Perbandingan Klasifikasi Berita Hoax Kategori Kesehatan Menggunakan Naïve Bayes dan Multinomial Naïve Bayes”* menggunakan 200 data berita dalam bidang kesehatan dari beberapa laman *website*, dengan tujuan untuk mengklasifikasi ke dalam kelas *hoax* dan *non-hoax* dengan hasil akurasi 90% pada partisi 90% *data training* dan 10% *data testing*

untuk metode Multinomial *Naïve Bayes*. Penelitian lainnya juga dilakukan oleh Jasman Pardede, dkk., (2020) dengan judul “Deteksi Komentar *Cyberbullying* Pada Media Sosial Berbahasa Inggris Menggunakan *Naïve Bayes Classification*” dengan tujuan untuk mengklasifikasi data ke dalam kelas *bully* dan *non-bully* dengan menggunakan 2.736 data komentar dari Kaggle, pada partisi data 80% *data training* dan 20% *data testing* didapatkan hasil akurasi yaitu sebesar 80%.

Penelitian terdahulu lainnya oleh Reza El Akbar, dkk., (2019) dengan judul “*The Implementation of Naïve Bayes Algorithm for Classifying Tweets Containing Hate Speech with Political Motive*” mengklasifikasi 2.271 data *tweet* ke dalam tiga kelas yaitu *non-political hate speech*, *motivated speech politics*, dan *non-hate speech*, di mana penelitian ini mendapatkan nilai akurasi 94% pada partisi *data training* 80% dan *data testing* 20% dengan metode *Naïve Bayes*. Penelitian lain dilakukan oleh Dzadda Ulhaq, dkk., (2022) yang berjudul “Klasifikasi Berita Kriminal Menggunakan Algoritma *Naïve Bayes* Berbasis PSO” menggunakan 120 data berita yang didapatkan dari laman jpn.com/kriminal, dengan tujuan mengklasifikasi data ke dalam kelas berita narkoba dan berita pembunuhan dengan mendapatkan nilai akurasi 81,67%. Penelitian selanjutnya dengan judul “*Classification of Bullying Comments on Youtube Streamer Comment Section Using Naïve Bayes Classification*” oleh Ahlida Nikmatul, dkk., (2023) bertujuan untuk mengklasifikasi data komentar dari YouTube ke dalam kelas komentar positif (*bullying*) dan komentar negatif (*non-bullying*) dengan metode *Naïve Bayes* menghasilkan nilai akurasi sebesar 84% pada partisi *data training* 90% dan *data testing* 10%. Penelitian lainnya oleh Sunardi, dkk., (2023) yang berjudul “*Comparing Data Mining Classification for Online Fraud Victim Profile in*

Indonesia” dengan melakukan komparasi dari tiga metode yaitu *Naïve Bayes*, *Decision Tree*, dan *Random Forest* dalam mengklasifikasi profil pelaku kejahatan ke dalam kelas positif dan kelas negatif menggunakan 1.587 data dari narasumber, kemudian menghasilkan nilai akurasi 77,3 % dengan metode *Naïve Bayes* dan *Decision Tree*, sedangkan nilai akurasi 76,8% dengan metode *Random Forest*.

Berdasarkan latar belakang penelitian tersebut, maka dilakukan penelitian terkait fenomena atau kasus *Cyber Harassment* dengan menggunakan data unggahan berbahasa Indonesia di media sosial Twitter dengan metode Multinomial *Naïve Bayes*. Pada penelitian terdahulu yang telah disebutkan pada poin sebelumnya, tidak ada penelitian yang mengklasifikasikan ke dalam empat kelas dan menggunakan topik spesifik *Cyber Harassment*. Penelitian ini memiliki *gap* atau perbedaan dengan penelitian-penelitian sebelumnya, di mana penelitian ini bertujuan untuk mengklasifikasi data ke dalam empat jenis *Cyber Harassment* yaitu kelas *Physical Threats*, *Purposeful Embarrassment*, *Racist*, dan *Sexual Harassment*. Penelitian ini diharapkan dapat menentukan jenis atau kelas *Cyber Harassment* yang tepat pada data yang diberikan menggunakan metode Multinomial *Naïve Bayes*.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan di atas, dapat dirumuskan permasalahan dalam penelitian ini yaitu:

1. Bagaimana hasil klasifikasi jenis-jenis *Cyber Harassment* di media sosial Twitter menggunakan Multinomial *Naïve Bayes*?
2. Bagaimana tingkat akurasi dari klasifikasi jenis-jenis *Cyber Harassment* di media sosial Twitter menggunakan Multinomial *Naïve Bayes*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah tersebut, maka penelitian ini dibuat dengan tujuan sebagai berikut:

1. Mengetahui hasil klasifikasi jenis-jenis *Cyber Harassment* di media sosial Twitter menggunakan Multinomial *Naïve Bayes*.
2. Mengetahui tingkat akurasi dari klasifikasi jenis-jenis *Cyber Harassment* di media sosial Twitter menggunakan Multinomial *Naïve Bayes*.

1.4 Manfaat Penelitian

Berdasarkan rumusan masalah dan tujuan penelitian, penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Memberikan informasi terkait performa klasifikasi dari metode Multinomial *Naïve Bayes*.
2. Memberikan informasi terkait kata-kata yang merepresentasikan perilaku *Cyber Harassment* di media sosial Twitter, sehingga dapat digunakan sebagai gambaran maupun bahan evaluasi untuk meminimalisir tingkat kejahatan *Cyber Harassment*.
3. Memberikan referensi dalam mengkaji dan menganalisis terkait pengklasifikasian data teks dengan menggunakan metode Multinomial *Naïve Bayes*.

1.5 Batasan Masalah

Batasan masalah pada penelitian ini sebagai berikut:

1. Data penelitian adalah data *tweet Cyber Harassment* tahun 2023 berbahasa Indonesia.

2. *Tweet* akan diklasifikasikan berdasarkan hasil penelitian yang dilakukan oleh Pew Research Center ke dalam empat kelas jenis-jenis *Cyber Harassment* yaitu *Physical Threats*, *Purposeful Embarrassment*, *Racist*, dan *Sexual Harassment*.

BAB II

TINJAUAN PUSTAKA

2.1 Teorema Bayes

Teorema *Bayes* merupakan teorema yang mengacu pada konsep probabilitas bersyarat, yakni probabilitas terjadinya kejadian A apabila diketahui bahwa suatu kejadian lain B telah terjadi (Walpole, 1995). Teorema *Bayes* diperkenalkan oleh Thomas *Bayes* (1701-1761) yang kemudian dikembangkan oleh Pierre-Simon Laplace (1749-1827). Teorema *Bayes* merupakan teorema yang menjelaskan suatu kemungkinan atau probabilitas dari sebuah hipotesis dapat berubah ketika mendapatkan bukti atau informasi baru (Stone, 2013).

Misalkan dalam suatu kejadian terdapat dua kemungkinan peristiwa yang terjadi, yaitu peristiwa A dan peristiwa B , yang mana kedua peristiwa tersebut saling berhubungan (dependen). Maka probabilitas bersyarat peristiwa A , apabila peristiwa B telah terjadi didefinisikan sebagai berikut (Murphy, 2012):

$$P(A|B) = \frac{P(A \cap B)}{P(B)} ; P(B) > 0 \quad (2.1)$$

Dengan $P(A \cap B)$ merupakan probabilitas kejadian gabungan A dan B yang didefinisikan sebagai $P(A|B)P(B)$, dan $P(B)$ adalah probabilitas B . Sedangkan probabilitas bersyarat peristiwa B , apabila peristiwa A telah terjadi didefinisikan sebagai berikut:

$$P(B|A) = \frac{P(B \cap A)}{P(A)} ; P(A) > 0 \quad (2.2)$$

Dengan $P(B \cap A)$ merupakan probabilitas kejadian gabungan B dan A yang didefinisikan sebagai $P(B|A)P(A)$ dan $P(A)$ adalah probabilitas A .

Berdasarkan teori himpunan, diketahui kejadian $A \cap B$ dan $B \cap A$ setara, sehingga didapatkan Teorema *Bayes* pada persamaan (2.3).

$$\begin{aligned} P(A \cap B) &= P(B \cap A) \\ P(A|B)P(B) &= P(B|A)P(A) \\ P(A|B) &= \frac{P(B|A)P(A)}{P(B)} \end{aligned} \quad (2.3)$$

Pada persamaan (2.3), variabel A menyatakan suatu hipotesis dan variabel B merupakan bukti atau *evidence*, sehingga $P(A|B)$ merupakan probabilitas hipotesis A berdasarkan bukti B yang telah diamati. Teorema *Bayes* menggunakan dasar pemikiran bahwa pengetahuan akan probabilitas sebelumnya (*prior*) dapat mempengaruhi probabilitas setelahnya (*posterior*), atau dapat didefinisikan sebagai teorema yang melakukan prediksi kemungkinan di masa depan dapat terjadi berdasarkan petunjuk atau informasi baru pada masa sebelumnya (Zhai dan Massung, 2016).

2.1.1 *Naïve Bayes Classifier*

Naïve Bayes Classifier (NBC) merupakan algoritma yang dipergunakan untuk mencari nilai probabilitas tertinggi pada data uji atau *data testing* untuk mengklasifikasi pada kelas atau kategori yang tepat (Feldman dan Sanger, 2007). NBC merupakan metode yang digunakan untuk pengklasifikasian data berdasarkan Teorema *Bayes* dengan mengasumsikan data yang digunakan mempunyai sifat tidak saling terikat antar satu sama lain (*independen*). Metode ini disebut “*naïve*” dikarenakan asumsi independensi yang tinggi (Roifa, 2018). Menggunakan persamaan (2.3), algoritma NBC adalah sebagai berikut:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} \quad (2.4)$$

Pada persamaan (2.4), variabel Y merepresentasikan kelas, sedangkan variabel X merepresentasikan atribut atau prediktor. Dalam proses pengklasifikasiannya, NBC membutuhkan sejumlah arahan yaitu dari prediktor-prediktor untuk mendapatkan kelas yang paling tepat (Rinaldi, dkk., 2021), sehingga persamaan (2.4) disesuaikan sebagai berikut:

$$P(Y|X_1, X_2, \dots, X_n) = \frac{P(Y)P(X_1, X_2, \dots, X_n|Y)}{P(X_1, X_2, \dots, X_n)} \quad (2.5)$$

Persamaan (2.5) dapat dituliskan secara sederhana sebagai berikut:

$$Posterior = \frac{Prior \times Likelihood}{Evidence} \quad (2.6)$$

Mengklasifikasikan data pada sebuah kelas dapat ditentukan dengan memilih kelas yang mempunyai nilai *posterior* tertinggi, di mana untuk mendapatkan nilai *posterior* yaitu dengan cara mengalikan probabilitas kemunculan kelas (*prior*) dengan probabilitas kemunculan atribut atau prediktor pada kelas tersebut (*likelihood*). Sedangkan diketahui dalam klasifikasi metode NBC, *evidence* untuk setiap kelas akan bernilai konstan, sehingga nilai *evidence* dapat diabaikan (Rinaldi, dkk., 2021). Maka persamaan (2.5) dapat ditulis sebagai:

$$P(Y|X_1, X_2, \dots, X_n) = P(Y)P(X_1, X_2, \dots, X_n|Y) \quad (2.7)$$

Probabilitas *posterior* $P(Y|X_1, X_2, \dots, X_n)$ selanjutnya dapat dijabarkan menggunakan aturan perkalian berdasarkan definisi peluang bersyarat pada persamaan (2.7) menjadi sebagai berikut:

$$\begin{aligned} P(Y|X_1, \dots, X_n) &= P(Y)P(X_1, X_2, \dots, X_n|Y) \\ &= P(Y)P(X_1|Y)P(X_2, X_3, \dots, X_n|Y, X_1) \\ &= P(Y)P(X_1|Y)P(X_2|Y, X_1)P(X_3, X_4, \dots, X_n|Y, X_1, X_2) \\ &= P(Y)P(X_1|Y)P(X_2|Y, X_1) \dots P(X_n|Y, X_1, X_2, \dots, X_{n-1}) \end{aligned} \quad (2.8)$$

Hasil penjabaran persamaan (2.8) diketahui semakin banyak dan kompleksnya faktor-faktor syarat yang dapat mempengaruhi nilai probabilitas yang hampir mustahil untuk dianalisis satu persatu. Kompleksitas dari faktor-faktor tersebut menyebabkan adanya kesulitan untuk melakukan perhitungan terhadap nilai probabilitas, maka dari itu digunakan asumsi independensi yang sangat tinggi (*naïve*). Asumsi tersebut menganggap bahwa setiap atribut atau prediktor X_1, X_2, \dots, X_n adalah saling bebas (independen) satu sama lain sehingga berlaku suatu persamaan sebagai berikut:

$$\begin{aligned} P(X_i|X_j) &= \frac{P(X_i \cap X_j)}{P(X_j)} \\ &= \frac{P(X_i)P(X_j)}{P(X_j)} \\ &= P(X_i) \end{aligned} \quad (2.9)$$

Persamaan (2.9) menunjukkan asumsi independensi yang tinggi yang dimiliki oleh metode NBC yang membuat syarat peluang atau probabilitas menjadi sangat sederhana, sehingga perhitungan menjadi mungkin untuk dilakukan. Maka persamaan probabilitas *posterior* $P(Y|X_1, X_2, \dots, X_n)$ dapat disederhanakan menjadi:

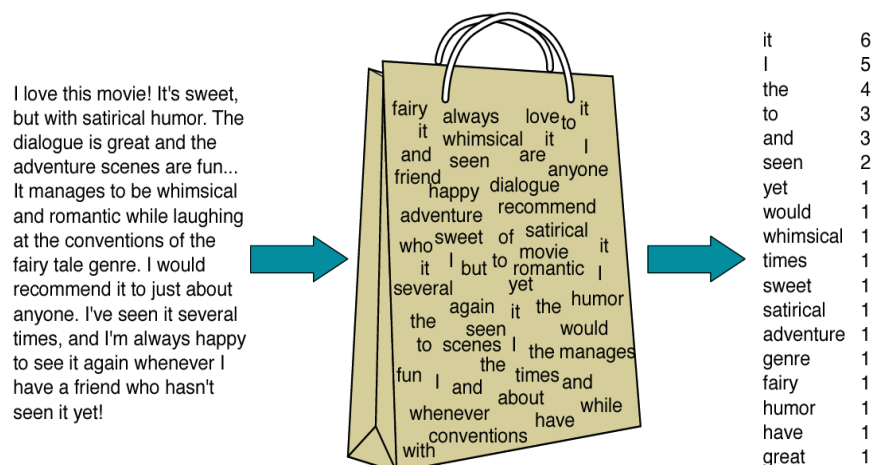
$$\begin{aligned} P(Y|X_1, X_2, \dots, X_n) &= P(Y)P(X_1|Y)P(X_2|Y)P(X_3|Y) \dots P(X_n|Y) \\ &= P(Y) \prod_{i=1}^n P(X_i|Y) \end{aligned} \quad (2.10)$$

2.1.2 Multinomial Naïve Bayes

Multinomial *Naïve Bayes* (MNB) merupakan salah satu jenis dari metode *Naïve Bayes Classifier* yang sering digunakan untuk membangun model pada kasus klasifikasi teks dalam bidang *machine learning*. *Machine learning* atau pembelajaran mesin adalah suatu praktik komputer *programming* yang digunakan untuk mempelajari data. *Machine learning* memerlukan data sebagai

bahan pembelajaran untuk menemukan pola atau *pattern* data sehingga dapat diselesaikan dan mendapatkan informasi yang berguna (Nurhafidzah, 2019). *Machine learning* memiliki beberapa jenis sistem pembelajaran, di mana untuk permasalahan klasifikasi termasuk ke dalam *supervised learning* (Russell, 2018). *Supervised learning* dalam *machine learning* adalah pembelajaran yang menggunakan data latih atau label sebagai satu set contoh yang akan digunakan untuk mengklasifikasi data (Garcia, dkk., 2015)

Multinomial *Naïve Bayes* (MNB) merupakan metode yang akan menghitung probabilitas bersyarat dari sebuah eksperimen dengan mengalikan probabilitas prior dengan *likelihood* pada setiap percobaan. Multinomial *Naïve Bayes* memodelkan distribusi kata dalam sebuah dokumen berdasarkan distribusi multinomial, di mana sebuah dokumen diperlakukan sebagai barisan kata dan diasumsikan bahwa setiap posisi untuk setiap kata adalah independen atau tidak mempengaruhi satu sama lain (Rennie, dkk., 2003). Metode MNB juga dikenal sebagai “*bag of words assumption*” atau metode yang memiliki asumsi pada kantung kata yang tidak mementingkan urutan (Reiten, 2017).



Gambar 2.1 *Bag of Words Assumption* pada Multinomial *Naïve Bayes*

Dalam klasifikasi teks, dokumen D yang telah diberi label direpresentasikan dengan $D = \{W_1, W_2, \dots, W_n, C\}$, dengan variabel atau vektor W merepresentasikan sebuah kata pada dokumen D , dan C adalah label kelas dari dokumen D . Menggunakan algoritma NBC pada persamaan (2.10), diketahui algoritma MNB adalah sebagai berikut (Schutze, dkk., 2008):

$$P(C|D) = P(C) \prod_{i=1}^n P(W_i|C) \quad (2.11)$$

dengan

$P(C|D)$: Probabilitas *posterior* kelas C benar terjadi dengan diberikan dokumen D

$P(C)$: Probabilitas *prior* dari kelas C

$P(W_i|C)$: Probabilitas kata ke- i pada dokumen dengan diberikan kelas C (*likelihood*)

W_1, W_2, \dots, W_n : Kata ke-1, 2, ..., n dalam dokumen

Pada persamaan (2.11), variabel C merepresentasikan kelas, sedangkan variabel D merepresentasikan dokumen. Sehingga dapat diinterpretasikan bahwa $P(W_i|C)$ dapat mengukur seberapa banyak bukti yaitu kemunculan sebuah kata W dalam dokumen yang memberikan kontribusi bahwa C adalah kelas yang tepat (Schutze, dkk., 2008). Nilai probabilitas *prior* $P(C)$ dapat dihitung menggunakan persamaan (2.12) sebagai berikut:

$$P(C = c_i) = \frac{n_{c_i}}{N_d} \quad (2.12)$$

dengan

c_i : Sub-kelas c ke- i

n_{c_i} : Jumlah kelas c ke- i

N_d : Total keseluruhan kelas

Untuk menghitung nilai *likelihood*, dilakukan pendekatan dengan menggunakan distribusi multinomial. Misalkan $x_c = \{x_{c_1}, x_{c_2}, \dots, x_{c_n}\}$ merupakan variabel acak diskrit yang merupakan vektor untuk menghitung jumlah kemunculan kata W yang muncul dengan total N dalam dokumen, maka untuk memaksimalkan nilai *likelihood* adalah sebagai berikut (Jebreen, 2017):

$$\begin{aligned} P(W_i = x_{c_i} | C = c_i) &= \frac{N!}{x_{c_1}! x_{c_2}! \dots x_{c_n}!} \theta_{c_1}^{x_{c_1}} \theta_{c_2}^{x_{c_2}} \dots \theta_{c_n}^{x_{c_n}} \\ &= \frac{N!}{\prod_{i=1}^n x_{c_i}!} \prod_{i=1}^n \theta_{c_i}^{x_{c_i}} \end{aligned} \quad (2.13)$$

dengan

W_i	: Kata ke- i
x_{c_i}	: Jumlah kemunculan kata ke- i pada kelas c
C	: Kelas yang dicari
c_i	: Sub-kelas c ke- i yang dicari
N	: Jumlah seluruh kata yang muncul
θ_{c_i}	: Probabilitas kemunculan kata ke- i pada kelas c

dengan nilai $\sum_{i=1}^n \theta_{c_i} = 1$ dan $\sum_{i=1}^n x_{c_i} = N$.

Estimasi parameter dalam MNB dapat menggunakan *Maximum Likelihood Estimate* (MLE). Metode MLE dapat digunakan apabila didasarkan pada suatu distribusi tertentu, di mana pada MNB didasarkan pada distribusi multinomial. Dalam mencari nilai estimasi parameter menggunakan MLE, langkah pertama yang harus dilakukan adalah mendefinisikan fungsi *likelihood* dari *probability mass function* distribusi multinomial sebagai berikut:

$$L(\theta_{c_i}) = L(\theta_{c_1}, \dots, \theta_{c_n}) = C(x) \prod_{i=1}^n \theta_{c_i}^{x_{c_i}} \quad (2.14)$$

dengan $C(x) = \frac{N!}{x_{c_1}! x_{c_2}! \dots x_{c_n}!}$

Selanjutnya menentukan fungsi log-likelihood $\ln L(\theta_{c_i})$ yaitu sebagai berikut:

$$\begin{aligned}\ell(\theta_{c_1}, \dots, \theta_{c_n}) &= \log L(\theta_{c_1}, \dots, \theta_{c_n}) \\ &= \log C(x) + \log \prod_{i=1}^n \theta_{c_i}^{x_{c_i}} \\ &= \ln C(x) + \sum_{i=1}^n x_{c_i} \ln \theta_{c_i}\end{aligned}\quad (2.15)$$

Diketahui $\sum_{i=1}^n \theta_{c_i} = 1$ dan $\sum_{i=1}^n x_{c_i} = N$, maka $\theta_{c_n} = 1 - \sum_{i=1}^{n-1} \theta_{c_i}$ dan $x_{c_n} = N - \sum_{i=1}^{n-1} x_{c_i}$, sehingga kita memiliki:

$$\begin{aligned}\ell(\theta_{c_1}, \dots, \theta_{c_{n-1}}) &= \ln L(\theta_{c_1}, \dots, \theta_{c_{n-1}}) \\ &= x_{c_1} \ln \theta_{c_1} + \dots + x_{c_i} \ln \theta_{c_i} + \dots + x_{c_{n-1}} \ln \theta_{c_{n-1}} \\ &\quad + x_{c_n} \ln \left(1 - \sum_{i=1}^{n-1} \theta_{c_i}\right)\end{aligned}\quad (2.16)$$

Untuk memaksimumkan fungsi *likelihood*, maka dicari persamaan diferensial dari $\ln L(\theta_{c_i})$ adalah:

$$\frac{\partial \ell}{\partial \theta_{c_i}} = \frac{\partial \ell(\theta_{c_1}, \dots, \theta_{c_{n-1}})}{\partial \theta_{c_i}} = \frac{x_{c_i}}{\theta_{c_i}} - \frac{x_{c_n}}{\theta_{c_n}} = 0, \quad i = 1, \dots, n-1$$

Sehingga,

$$\frac{x_{c_i}}{\theta_{c_i}} = \frac{x_{c_n}}{\theta_{c_n}}\quad (2.17)$$

Oleh karena itu, persamaan (2.17) memenuhi persamaan:

$$\frac{x_{c_1}}{\theta_{c_1}} = \frac{x_{c_n}}{\theta_{c_n}}, \frac{x_{c_2}}{\theta_{c_2}} = \frac{x_{c_n}}{\theta_{c_n}}, \dots, \frac{x_{c_{n-1}}}{\theta_{c_{n-1}}} = \frac{x_{c_n}}{\theta_{c_n}}\quad (2.18)$$

Persamaan (2.18) ekuivalen dengan:

$$\frac{x_{c_1}}{x_{c_n}} = \frac{\theta_{c_1}}{\theta_{c_n}}, \frac{x_{c_2}}{x_{c_n}} = \frac{\theta_{c_2}}{\theta_{c_n}}, \dots, \frac{x_{c_{n-1}}}{x_{c_n}} = \frac{\theta_{c_{n-1}}}{\theta_{c_n}}\quad (2.19)$$

$$\text{Menggunakan} \quad \sum_{i=1}^n \theta_{c_i} = 1 = \sum_{i=1}^n \frac{x_{c_i}}{x_{c_n}} \theta_{c_n} = \frac{\theta_{c_n}}{x_{c_n}} \sum_{i=1}^n x_{c_i} = N \frac{\theta_{c_n}}{x_{c_n}},$$

sehingga nilai MLE untuk $\theta_{c_1}, \dots, \theta_{c_n}$ adalah sebagai berikut:

$$\widehat{\theta}_{c_1} = \frac{x_{c_1}}{N}, \widehat{\theta}_{c_2} = \frac{x_{c_2}}{N}, \dots, \widehat{\theta}_{c_n} = \frac{x_{c_n}}{N} \quad (2.20)$$

Maka nilai estimasi parameter θ_{c_i} menggunakan MLE adalah:

$$\widehat{\theta}_{c_i} = \frac{x_{c_i}}{N} \quad (2.21)$$

Selanjutnya, untuk mengatasi probabilitas nol (*zero probability*), dilakukan cara sederhana yaitu dengan menggunakan teknik *additive smoothing*. *Additive smoothing* (atau *Laplace smoothing*) adalah metode yang digunakan untuk pemulusan (*smoothing*) atau untuk mengatasi nilai peluang nol. *Additive smoothing* dalam *Naïve Bayes* umumnya menambahkan nilai $\alpha = 1$ pada pembilang dan α_d yang merepresentasikan jumlah *vocabulary* dalam data teks pada penyebut (Schutze, 2008). Sehingga untuk menghitung nilai *likelihood* $P(W_i|C)$ pada MNB menggunakan MLE dengan *smoothing* dapat menggunakan persamaan (2.22):

$$\widehat{\theta}_{c_i} = \frac{x_{c_i} + \alpha}{N + \alpha_d} \quad (2.22)$$

dengan

- x_{c_i} : Jumlah kemunculan kata ke- i pada kelas c
- N : Jumlah seluruh kata yang muncul pada kelas yang diketahui
- α : *Add-one smoothing* untuk setiap kata pada kelas c
- α_d : Total seluruh kata berbeda pada seluruh dokumen

Tahapan yang dilalui dalam proses klasifikasi data teks dengan metode

Multinomial *Naïve Bayes* adalah sebagai berikut (Awan dan Naviani, 2023):

1. Menghitung probabilitas *prior* pada data, yaitu probabilitas setiap kelas dengan persamaan (2.12).
2. Menghitung nilai *likelihood* pada data, yaitu probabilitas setiap fitur (kata) pada masing-masing kelas dengan persamaan (2.22).

3. Menghitung probabilitas *posterior* dengan mengalikan probabilitas *prior* dengan *likelihood* dengan persamaan (2.11).
4. Mencari nilai probabilitas *posterior* tertinggi untuk menentukan kelas yang terpilih.
5. Menghitung ketepatan hasil klasifikasi dengan menghitung hasil klasifikasi yang tepat dan hasil klasifikasi yang tidak tepat.

Tabel 2.1 Tabulasi Silang Klasifikasi *Multiclass*

		<i>Predicted class</i>			
		<i>Physical Threats</i>	<i>Purposeful Embarrassment</i>	<i>Racist</i>	<i>Sexual Harassment</i>
<i>Actual class</i>	<i>Physical Threats</i>	T_{11}	F_{12}	F_{13}	F_{14}
	<i>Purposeful Embarrassment</i>	F_{21}	T_{22}	F_{23}	F_{24}
	<i>Racist</i>	F_{31}	F_{32}	T_{33}	F_{34}
	<i>Sexual Harassment</i>	F_{41}	F_{42}	F_{43}	T_{44}

Pada tabel (2.1) ditunjukkan tabulasi silang untuk klasifikasi *multiclass* dengan $T_{mn} = T_{11}$ merepresentasikan *true class* pada baris ke- $m = 1$ dan kolom ke- $n = 1$ dan $F_{mn} = F_{12}$ merepresentasikan *false class* pada baris ke- $m = 1$ dan kolom ke- $n = 2$. Sehingga dapat dihitung tingkat akurasi klasifikasi MNB dengan menggunakan *accuracy*. *Accuracy* merupakan rasio perbandingan mengenai informasi yang diprediksi benar oleh model dari keseluruhan informasi, yang dapat diformulakan sebagai berikut (Roifa, 2018):

$$\begin{aligned}
 Accuracy &= \frac{\text{Jumlah prediksi benar}}{\text{Jumlah keseluruhan prediksi}} \\
 &= \frac{T_{11} + T_{22} + T_{33} + T_{44}}{T_{11} + F_{12} + F_{13} + F_{14} + F_{21} + T_{22} + F_{23} + F_{24} + F_{31} + F_{32} + T_{33} + F_{34} + F_{41} + F_{42} + F_{43} + T_{44}}
 \end{aligned}
 \tag{2.23}$$

Tabel 2.2 menunjukkan interval dari persentase nilai akurasi klasifikasi yang dapat menjadi acuan dalam menginterpretasi seberapa baik model dalam mengklasifikasi data (Stephen, 2022):

Tabel 2.2 Interval Kategori Nilai Akurasi

Interval	Kategori
91% – 100%	Sangat Baik
71% – 90%	Baik
61% – 70%	Cukup Baik
< 60%	Kurang Baik

Sedangkan dapat dihitung pula nilai eror atau *Apparent Error Rate* (APER) yang akan menyatakan seberapa besar kesalahan pada klasifikasi MNB. Nilai eror atau *misclassification* pada umumnya berkisar antara 0 – 0,5 atau dalam presentase < 50%. APER merupakan rasio perbandingan mengenai informasi yang salah diprediksi oleh model dari keseluruhan informasi, yang dapat diformulakan sebagai berikut (Rinaldi, 2021):

$$\begin{aligned}
 APER &= \frac{\text{Jumlah prediksi salah}}{\text{Jumlah keseluruhan prediksi}} \\
 &= \frac{F_{12}+F_{13}+F_{14}+F_{21}+F_{23}+F_{24}+F_{31}+F_{32}+F_{34}+F_{41}+F_{42}+F_{43}}{T_{11}+F_{12}+F_{13}+F_{14}+F_{21}+T_{22}+F_{23}+F_{24}+F_{31}+F_{32}+T_{33}+F_{34}+F_{41}+F_{42}+F_{43}+T_{44}}
 \end{aligned}
 \tag{2.24}$$

2.2 Twitter

Twitter merupakan media sosial jenis *microblogging* yang dioperasikan oleh Twitter Inc., yang memungkinkan penggunaanya untuk bertukar pesan (Twitter, 2016). Mikroblog merupakan sebuah alat komunikasi yang digunakan secara *online* yang memberi akses kepada penggunaanya untuk memperbarui status atau informasi mengenai kehidupannya yang hendak dibagikan. Twitter menjadi media sosial yang membuat penggunaanya dapat mengetahui topik yang sedang *trending* dengan fitur *trending topic* dan dapat memudahkan penggunaanya untuk mencari topik, nama,

atau informasi tertentu dengan lebih spesifik melalui kolom *search* (Kurniawan, 2017).

Twitter menyediakan *Application Programming Interface* (API) yang dapat digunakan untuk mengumpulkan atau mengolah data dengan menyediakan kode atau *key*. Twitter API memungkinkan *developer* untuk memanggil fungsi tertentu pada pengembangan perangkat lunak yang sedang dibangun (Rahmi, 2021). Dengan menggunakan Twitter API, pengguna dapat mengumpulkan data menggunakan kata kunci tertentu dengan bantuan bahasa pemrograman seperti *R* dan *Python*.

2.3 Cyber Harassment

Cyber harassment atau *cyber abuse* atau pelecehan *online* didefinisikan sebagai tindakan yang memanfaatkan beragam jenis teknologi komputer untuk melakukan kejahatan *online* seperti perundungan, menguntit, dan pelecehan seksual. *Cyber harassment* merupakan penggunaan internet, *email*, atau sejenisnya yang bertujuan untuk mengganggu, mengancam, dan membahayakan orang lain (Winkelman, dkk., 2015). *Cyber harassment* atau *cyber abuse* merupakan penggunaan segala bentuk media digital termasuk *chatroom*, *email*, situs web, ponsel dan sebagainya yang dapat digunakan sebagai anonim untuk mempermalukan, melecehkan, dan mengancam korban yang dapat menyebabkan tekanan emosional (Wolford, 2016). Jenis tindakan atau media *cyber harassment* di antaranya sebagai berikut:

1. Mengirim pesan yang mengancam atau melecehkan.
2. Menyebarkan data pribadi korban atau mempergunakannya untuk tindak kejahatan.

3. Menggunakan internet untuk mempermalukan atau melecehkan korban.
4. Mengunggah konten yang tidak pantas dan tanpa persetujuan di media sosial.

Pew Research Center merupakan salah satu pusat penelitian non-pemerintah di Amerika Serikat, yang fokus penelitiannya pada bidang masalah sosial, opini publik, dan tren demografis. Dalam penelitiannya yang berjudul “*The State of Online Harassment*” (2021), Pew Research Center memaparkan bahwa 41% orang Amerika mengalami *cyber harassment*, di mana jenis *cyber harassment* tersebut adalah ancaman kekerasan fisik (*physical threats*), mengejek/mempermalukan (*purposeful embarrassment*), panggilan yang menyinggung (*racist*), dan pelecehan seksual (*sexual harassment*). Hasil penelitian tersebut mengungkapkan bahwa faktor yang melatarbelakangi pelecehan *online* adalah karena pandangan politik (50%), jenis kelamin (33%), ras atau etnis (29%), agama (19%), dan orientasi seksual (16%).

2.3.1 *Physical Threats*

Physical threats (kekerasan fisik) didefinisikan sebagai tindakan yang mengakibatkan atau memberi dampak gangguan baik pada fisik maupun mental. Kekerasan merupakan perbuatan yang membuat seseorang cedera atau terluka, atau kerusakan yang terjadi pada barang (Diyah, 2016). Suyanto (Putri, 2022) membagi beberapa jenis kekerasan fisik dalam beberapa kategori yakni memukul, mendorong, menampar, mengancam, menendang, dan sebagainya. Sedangkan ancaman terhadap kekerasan fisik merupakan perilaku mengancam yang melibatkan fisik, yang dapat memberi dampak pada kestabilan mental korban. Ancaman-ancaman yang mengarah pada kekerasan fisik di media sosial biasanya mengandung kata-kata yang merupakan bentuk dari kekerasan fisik itu

sendiri, di antaranya adalah “bunuh”, “pukul”, “tampar”, “jambak”, dan “tendang” (Kango, 2009).

2.3.2 Purposeful Embarrassment

Purposeful embarrassment (mengejek atau mengolok) merupakan perbuatan yang secara sengaja dilakukan dan mengakibatkan ketersinggungan dari pihak yang dirugikan. Perilaku mengejek di media sosial dapat mengarah pada penghinaan atau pencemaran nama baik (Kusno, 2021). Mengejek juga dapat disebut sebagai verbal *bullying* (Putri, dkk., 2021) yang didefinisikan sebagai bentuk penindasan atau penghinaan dengan memakai kata-kata yang tidak pantas dengan maksud mengolok, mengejek, dan berkata kasar. Kata-kata negatif yang biasanya digunakan untuk mengejek atau mengolok-olok mengandung unsur nama hewan, menyinggung fisik, dan nama makhluk halus. Sedangkan kata-kata yang paling sering muncul dalam perilaku mengejek di media sosial di antaranya adalah “bego”, “asu”, “sampah” (Rakhmawati, dkk., 2021) dan “*coward*” (pecundang), “*ugly*” (jelek), “*idiot*” (Gorell, dkk., 2020).

2.3.3 Racist

Racist (rasis) adalah perbuatan atau sikap yang memiliki upaya untuk menciptakan kebencian, kekerasan, dan diskriminasi baik terhadap individu atau kelompok berdasarkan faktor warna kulit, kewarganegaraan, suku, dan agama (Amin, dkk., 2022). Rasisme adalah sebuah gagasan bahwa manusia terbagi ke dalam sejumlah ras dan suatu ras dianggap memiliki derajat yang lebih rendah dibanding ras lainnya (Tirahmawan, dkk, 2021). Tindakan *racist* di media sosial menjadi salah satu ujaran kebencian (*hate speech*) yang umumnya ditemukan dengan menyinggung masalah ras, suku, agama, atau golongan yang

saling membenci atau merendahkan (Amin, dkk., 2018). Kata yang sering ditemukan untuk tindakan *racist* di media sosial adalah “cina”, “kafir”, dan “Islam” (Ash-Shidiq dan Pratama. 2021). Sedangkan penelitian mengenai *hate speech* di media sosial (Ibrahim, dkk., 2021) didapatkan bahwa kata-kata yang sering digunakan sejumlah oknum pelaku *racist* di media sosial adalah “muslim”, “*mosque*” (masjid), “*jewish*” (yahudi), “*black*” (hitam), dan “*white*” (putih).

2.3.4 Sexual Harassment

Sexual harassment (pelecehan seksual) adalah salah satu bentuk dari kekerasan yang tindakannya adalah memberi atensi secara tidak lazim dan tidak diinginkan secara seksual. Dalam ranah hukum, “*sexual harassment*” diartikan sebagai “tindak pemaksaan suatu kehendak seksual atau penyerangan secara seksual” (Mustika, dkk., 2021). Pelecehan seksual di media sosial ditemukan dalam bentuk kata-kata atau gambar dengan sasaran korban dalam berbagai rentang usia (Kurniati, 2022). Perilaku pelecehan seksual di media sosial biasanya ditemukan dalam bentuk kata-kata atau kalimat, di mana kata yang paling banyak ditemukan dalam kasus pelecehan seksual adalah “perkosa”, “sentuh”, “raba”, “*cat calling*”, dan “pelecehan seksual” (Putri, dkk., 2020). Sedangkan pada penelitian mengenai *cyber harassment* di sosial media (Abarna, dkk., 2022) diketahui kata yang sering digunakan oleh pelaku pelecehan seksual adalah “*bitch*” (jalang).

2.4 Data Crawling

Teknik *data crawling* merupakan sebuah cara yang dapat digunakan untuk mengumpulkan data penelitian. *Data crawling* dapat mengumpulkan data atau

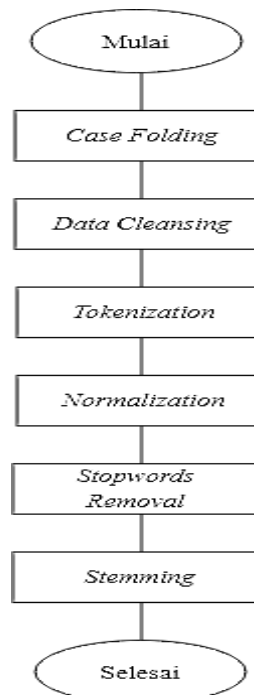
informasi yang berada di halaman *web* atau media sosial dengan cara mengunduh data dari suatu *database*. Teknik *data crawling* akan bekerja secara otomatis berdasarkan pada kata kunci yang ditentukan oleh pengguna (Saputra, 2017). Pada Twitter, telah tersedia *Application Programming Interface* (API) yang merupakan sebuah program yang dapat memudahkan pengguna untuk mengakses data atau informasi yang berada di Twitter. Twitter API tersebut dapat dimanfaatkan untuk mengumpulkan data penelitian menggunakan teknik *data crawling* dengan mengakses laman Twitter *Developer* sehingga akan mendapatkan kode atau *key* yang menjadi sebuah syarat untuk mengunduh data penelitian (Sembodo, 2016).

2.5 Data Preprocessing

Data preprocessing merupakan salah satu langkah penting dalam bidang *machine learning* yang bertujuan untuk menyaring atau membersihkan data agar dapat menghindari data yang tidak konsisten atau kurang sempurna sehingga dapat meningkatkan akurasi klasifikasi data (Garcia, dkk., 2015). *Data preprocessing* digunakan untuk menghilangkan fitur yang tidak relevan dan redundan yang mungkin akan menyebabkan kebingungan atau dapat mempengaruhi akurasi suatu metode (Gorunescu, 2011). *Data preprocessing* dilakukan untuk mengubah data yang tidak terstruktur atau semi-terstruktur seperti teks, gambar, maupun audio, menjadi data yang terstruktur (Kurniawan, 2017). *Data preprocessing* memiliki enam tahap sebagai berikut:

1. *Case folding*, bertujuan untuk menyeragamkan huruf menjadi huruf kecil, misalnya “Saya adalah WNI” menjadi “saya adalah wni”.

2. *Data cleansing*, bertujuan untuk menghapus *link*, *emoticon*, dan atribut-atribut lain yang tidak dibutuhkan, seperti *username* (@), *hashtag* (#), tanda baca (? , !, /, dan lain sebagainya), dan lain sebagainya.
3. *Tokenization*, bertujuan mengubah kalimat menjadi kata-kata yang terpisah dalam dokumen, misalnya “mari saling bertoleransi” menjadi “mari, saling, ber, toleransi”.
4. *Normalization*, bertujuan untuk menyeragamkan variasi kata yang memiliki arti yang sama, misalnya kata “km”, “kmu”, dan “kamu”, akan diseragamkan menjadi “kamu”.
5. *Stopwords removal*, bertujuan untuk menghapus kata hubung (dan, namun, dan lain sebagainya), imbuhan (ke-, di-, -an, -kah, dan sebagainya), dan lain sebagainya yang tidak diperlukan.
6. *Stemming*, bertujuan untuk mengubah setiap kata ke dalam bentuk kata dasar atau kata baku, misalnya kata “detil” menjadi “detail”.



Gambar 2.2 Alur Data Preprocessing

2.6 Ekstraksi Fitur

Ekstraksi fitur adalah suatu proses untuk mengubah data yang tidak terstruktur (dalam bentuk teks) menjadi data yang terstruktur (data numerik). Proses ini penting untuk dilakukan karena data teks tidak dapat langsung diolah oleh algoritma *machine learning* (Faisal dan Nugrahadi, 2020). *Bag of Words* (BOW) merupakan salah satu teknik yang dapat digunakan untuk mengekstraksi fitur. BOW didefinisikan sebagai sebuah model yang dapat mempelajari kosakata dari seluruh dokumen yang ada, yang kemudian akan dimodelkan tiap dokumen untuk menghitung jumlah kemunculan setiap kata. BOW merepresentasikan teks atau kumpulan kata yang berada di sebuah kantung (*bag*), tanpa memperhatikan urutan kata dalam kantung tersebut (Putri dan Hendrowati, 2018).

Tabel 2.3 Contoh Kalimat Dalam Dokumen

Dokumen	Kalimat
1.	Ibu pelit banget kayak orang cina!
2.	Mantan pantas ditampar dan ditendang

Dari dokumen pada tabel 2.3, akan dihitung frekuensi kemunculan setiap kata sebagai berikut:

Tabel 2.4 Frekuensi Kemunculan Kata Dalam Dokumen

Dok.	banget	cina	dan	di	ibu	kayak	mantan	orang	pantas	pelit	tampar	tendang
1.	1	1	0	0	1	1	0	1	0	1	0	0
2.	0	0	1	2	0	0	1	0	1	0	1	1

2.7 Uji Performa Model

Salah satu metode yang dapat melakukan uji performa model atau validasi model terbaik adalah metode *k-fold cross validation*. Metode ini menguji keefektifan dari model yang dibentuk dengan *resampling* (penyusunan ulang) di mana data akan dibagi menjadi 2 bagian yakni data *training* dan data *testing*. *K-fold*

cross validation akan menghasilkan model yang tidak bias dikarenakan metode ini akan mengobservasi seluruh data, di mana setiap data mempunyai kesempatan yang sama untuk terpilih sebagai data *training* ataupun data *testing* (Widyaningsih, dkk., 2021).

K-fold cross validation akan membagi data menjadi k bagian (*folds*) yang sama besar, dengan nilai k yang ditentukan oleh peneliti. Setelah data dibagi menjadi k bagian yang sama besar, $k - 1$ bagian akan digunakan sebagai data training untuk melatih model dan 1 bagian yang tersisa digunakan sebagai data testing untuk model. Nilai k yang ditentukan tidak disarankan terlalu besar atau terlalu kecil, karena akan menghasilkan model yang bias ataupun *overfitting*. Nilai $k = 10$ merupakan nilai yang disarankan dan umumnya digunakan pada eksperimen terdahulu yang dapat menghasilkan model terbaik (Kuhn dan Johnson, 2013).

2.8 Cyber Harassment dalam Pandangan Islam

Cyber harassment merupakan satu dari sekian banyaknya jenis kejahatan, di mana kejahatan tersebut dilakukan di media sosial. Dalam pandangan ahli hukum fikih Islam, kejahatan dianggap sebagai tindakan-tindakan pidana yang diberi hukuman. Kejahatan tersebut mencakup kejahatan terhadap jiwa raga manusia (melukai anggota tubuh), kejahatan terhadap kehormatan (perbuatan zina), kejahatan yang membuat kerusakan di muka bumi, dan lain sebagainya (Rahmawati, 2022). Dalam penelitian ini, kasus *cyber harassment* mencakup tindakan melecehkan, mengancam, memermalukan, dan mengucilkan di media sosial. Di mana tindakan-tindakan tersebut adalah tindakan negatif yang berasal dari perilaku dan perkataan manusia kepada manusia lainnya. Dalam Al-Qur'an

surah al-Hujurat ayat 11, telah diperintahkan agar manusia tidak mempermalukan atau merendahkan manusia lain.

“Wahai orang-orang yang beriman, janganlah suatu kaum mengolok-olok kaum yang lain (karena) boleh jadi mereka (yang diolok-olokkan itu) lebih baik daripada mereka (yang mengolok-olok) dan jangan pula perempuan-perempuan (mengolok-olok) perempuan lain (karena) boleh jadi perempuan (yang diolok-olok itu) lebih baik daripada perempuan (yang mengolok-olok).”

Ayat tersebut menyampaikan mengenai larangan mengolok-olok orang lain, yang dalam penelitian ini termasuk ke dalam *cyber harassment* di mana tindak melecehkan tersebut adalah menghina atau menjatuhkan suatu golongan ras, agama, warna kulit, dan lain sebagainya. Mengolok-olok juga termasuk pada tindakan mempermalukan dengan sengaja atau mengejek, di mana seseorang merasa lebih baik atau lebih unggul dari orang lain yang dicelanya. Sedangkan bentuk pelecehan secara seksual juga dibahas dalam Al-Qur'an, sebagaimana Allah *subhanahu wa ta'ala* melarang tersebut dalam surah An-Nur ayat 33.

”... Dan janganlah kamu paksa hamba sahaya perempuanmu untuk melakukan pelacuran, sedang mereka sendiri menginginkan kesucian karena kamu hendak mencari keuntungan kehidupan duniawi. Dan barangsiapa yang memaksa mereka, maka sesungguhnya Allah adalah Maha Pengampun lagi Maha Penyayang (kepada mereka) sesudah mereka dipaksa itu.”

Ayat tersebut menyampaikan bahwa Allah *subhanahu wa ta'ala* melindungi orang-orang yang telah dipaksa melakukan pelacuran, atau dapat dikatakan sebagai para korban pelecehan seksual. Kasus *cyber harassment* yang mencakup tindakan rasisme, kekerasan fisik, pelecehan seksual, dan mengejek, merupakan tindak kejahatan yang jelas dilarang dalam Al-Qur'an, dan akan diberi hukuman yang seadil-adilnya. Islam sebagai agama yang mulia menjunjung tinggi tindakan yang baik (*ma'ruf*) dan melarang tindakan yang melanggar (*munkar*).

2.9 Kajian *Cyber Harassment* dengan Teori Pendukung

Cyber harassment (pelecehan *online*) atau dapat juga disebut sebagai *cyber abuse* merupakan salah satu jenis *cybercrime* yang didefinisikan sebagai suatu tindakan yang memanfaatkan beragam jenis teknologi untuk melakukan kejahatan *online* seperti perundungan, menguntit, hingga pelecehan seksual. *Cyber harassment* menjadi salah satu isu sosial secara global yang terus mengalami peningkatan. Pew Research Center dalam penelitiannya yang berjudul “*The State of Online Harassment*” (2021) memaparkan bahwa *cyber harassment* diklasifikasikan menjadi empat jenis, yaitu ancaman kekerasan fisik (*physical threats*), mengejek/mempermalukan (*purposeful embarrassment*), panggilan yang menyinggung (*racist*), dan pelecehan seksual (*sexual harassment*).

Direktorat Tindak Pidana Siber Bareksrim Polri melalui portal Patrolisiber mengungkapkan terdapat 15.152 laporan atau pengaduan mengenai kejahatan siber di Indonesia sepanjang tahun 2021, di mana 3.101 laporan di antaranya merupakan kasus pelecehan dengan pengancaman dan pencemaran nama baik. Southeast Asia Freedom of Expression Network (SAFE-net) turut menyatakan bahwa terjadi peningkatan kasus *cyber harassment* pada tahun 2020 sebesar sepuluh kali lipat dibandingkan tahun sebelumnya. Komisi Nasional Anti Kekerasan terhadap Perempuan (2022) memaparkan data berdasarkan survei yang dilakukan oleh Koalisi Ruang Publik Aman (KRPA) tahun 2021 bahwa *cyber harassment* paling banyak terjadi di media sosial (42%) dan aplikasi *chat* (33%).

Indonesia menempati peringkat kelima di dunia sebagai negara dengan masyarakat yang paling banyak menggunakan media sosial Twitter, dengan jumlah 18,54 juta pengguna (Rizaty, 2022). Sejalan dengan pemaparan pada poin

sebelumnya, penelitian yang berjudul “*Crime and its fear in social media*” yang dilakukan oleh Prieto, dkk., (2020) menyatakan bahwa salah satu media sosial dengan tingkat kejahatan siber tertinggi adalah Twitter. Pada media sosial Twitter, setidaknya terdapat lima *tweet* atau unggahan setiap menitnya yang mengandung unsur kejahatan.

Penelitian ini melakukan pengumpulan data dengan teknik *data crawling* di media sosial Twitter dengan memanfaatkan *Application Programming Interface* (API). Proses pengumpulan data dengan *data crawling* dilakukan secara berulang-ulang dikarenakan keterbatasan mengunduh data di media sosial Twitter dan banyaknya *tweet* yang terunduh tetapi bukan termasuk jenis-jenis perilaku *Cyber Harassment*, sehingga perlu dilakukan proses *filtering* untuk menyaring data yang akan digunakan dalam penelitian. *Tweet* yang dikumpulkan adalah *tweet* berbahasa Indonesia mengenai *cyber harassment* pada tahun 2023 yang mengacu pada 36 kata kunci yang didapatkan dari hasil penelitian terdahulu, sehingga didapatkan data *tweet* berjumlah 2.000 yang kemudian diklasifikasikan ke dalam empat kelas jenis *Cyber Harassment* yaitu *Physical Threats*, *Purposeful Embarrassment*, *Racist*, dan *Sexual Harassment*.

Sebelum dilakukan klasifikasi dengan metode *Multinomial Naïve Bayes*, perlu dilakukan *data preprocessing* sebagai langkah untuk memperjelas fitur dan membuang atribut-atribut yang tidak diperlukan. Tahap pertama pada *data preprocessing* adalah *case folding* yang merupakan tahap untuk menyeragamkan huruf menjadi huruf kecil; tahap kedua yaitu *data cleansing* yang merupakan tahap untuk menghapus *link*, *username*, *emoticon*, dan atribut-atribut lain yang tidak dibutuhkan; tahap ketiga yaitu tahap *tokenization* yang merupakan tahap untuk

mengubah kalimat menjadi kata-kata yang terpisah dalam dokumen; tahap keempat yaitu *normalization* yang merupakan tahap untuk menyeragamkan variasi kata yang memiliki arti yang sama; tahap kelima yaitu tahap *stopwords removal* yang merupakan tahap untuk menghapus kata hubung, imbuhan, dan lain sebagainya yang tidak diperlukan; dan tahap terakhir yaitu tahap *stemming* yang merupakan tahap mengubah setiap kata ke dalam bentuk kata dasarnya. Selanjutnya dilakukan ekstraksi fitur dengan metode *Bags of Word* untuk mentransformasi dari data teks menjadi data numerik. *Bag of Words* (BOW) merupakan teknik menghitung frekuensi kemunculan data dalam dokumen. BOW akan memberikan informasi seberapa banyak sebuah kata muncul untuk mendukung proses klasifikasi.

Setelah semua proses persiapan tersebut selesai, maka dilakukan pembuatan model klasifikasi dengan metode *Multinomial Naïve Bayes*. Tahap pertama adalah menghitung probabilitas *prior* pada data, yaitu probabilitas setiap kelas; tahap kedua yakni menghitung nilai *likelihood* pada data, yaitu probabilitas setiap fitur (kata) pada masing-masing kelas; tahap ketiga adalah menghitung probabilitas *posterior* dengan mengalikan probabilitas *prior* dengan *likelihood*; tahap ke-empat yakni mencari nilai probabilitas *posterior* tertinggi untuk menentukan kelas yang terpilih; dan tahap terakhir yaitu menghitung ketepatan hasil klasifikasi dengan menghitung hasil klasifikasi yang tepat dengan *accuracy* dan hasil klasifikasi yang tidak tepat dengan *Apparent Error Rate* (APER).

Setelah mendapatkan hasil klasifikasi, dilakukan uji atau evaluasi performa model menggunakan *k-fold cross validation*. *K-fold cross validation* akan membagi data menjadi k bagian (*folds*) yang sama besar, dengan nilai k yang ditentukan oleh peneliti. Setelah data dibagi menjadi k bagian yang sama besar, $k - 1$ bagian akan

digunakan sebagai data training untuk melatih model dan 1 bagian yang tersisa digunakan sebagai data testing untuk model. Nilai $k = 10$ merupakan nilai yang disarankan dan umumnya digunakan pada eksperimen terdahulu yang dapat menghasilkan model terbaik.

BAB III

METODE PENELITIAN

3.1 Pendekatan Penelitian

Penelitian ini menggunakan penelitian kuantitatif dengan pendekatan deskriptif dan studi literatur. Penelitian kuantitatif merupakan penelitian yang menggunakan populasi atau sampel tertentu dengan mengumpulkan data dan mengolah data yang bersifat statistik. Sedangkan pendekatan deskriptif adalah metode untuk menganalisis hasil penelitian dan studi literatur digunakan untuk mengkaji atau mempelajari teori dari sumber-sumber yang kredibel seperti buku, jurnal, dan lain sebagainya (Sugiyono, 2011).

3.2 Teknik Pengumpulan Data

Teknik pengumpulan data pada penelitian dilakukan dengan menggunakan teknik *data crawling*. Pengumpulan data dengan teknik *data crawling* dilakukan dengan bantuan pemrograman *R* dan memanfaatkan fasilitas *Application Programming Interface* (API) Twitter yang dilakukan secara berulang-ulang hingga mendapatkan 2.000 data atau *tweet* terkait *Cyber Harassment* dengan menggunakan kata kunci yang sudah ditentukan. Pengulangan pada proses *data crawling* dilakukan karena keterbatasan jumlah *tweet* yang dapat diunduh dan perlu dilakukannya proses *filtering* secara manual untuk menyaring atau memisahkan *tweet* yang bukan termasuk jenis-jenis *Cyber Harassment*. Kata kunci yang digunakan untuk mengumpulkan data didapatkan dari hasil studi literatur pada penelitian terdahulu yang ditunjukkan pada tabel 3.1.

Tabel 3.1 Kata Kunci

No.	Kelas	Kata Kunci
1.	<i>Physical Threats</i>	Bunuh, pukul, mukul, tampar, nampar, jambak, tendang, dan nendang.
2.	<i>Purposeful Embarrassment</i>	Bodoh, bego, tolol, goblok, asu, anjing, sampah, pecundang, pengecut, jelek, dan idiot.
3.	<i>Racist</i>	Islam, cina, muslim, masjid, yahudi, hitam, putih, dan kafir.
4.	<i>Sexual Harassment</i>	Perkosa, sentuh, raba, <i>cat calling</i> , pelecehan, leceh, jalang, lonte, dan pelacur.

3.3 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder berupa data unggahan atau *tweet* terkait kasus *Cyber Harassment* di media sosial Twitter yang dikumpulkan pada bulan Januari sampai dengan Juni tahun 2023. Jumlah data *tweet* dengan hasil pelabelan manual dapat dilihat pada tabel 3.2, sedangkan contoh *tweet* pada data penelitian dapat dilihat pada tabel 3.3 yang selengkapnya terlampir pada lampiran (1).

Tabel 3.2 Data Penelitian

No.	Kelas	Jumlah <i>Tweet</i>
1.	<i>Physical Threats</i>	500
2.	<i>Purposeful Embarrassment</i>	500
3.	<i>Racist</i>	500
4.	<i>Sexual Harassment</i>	500
Total		2.000

Tabel 3.3 *Tweet* dengan Pelabelan Kelas

No.	<i>User/ID</i>	<i>Tweet</i>	Kelas
1.	@tan*****	“nikah 3 kali dia main tangan??? <i>verbally abusive</i> ?? doyan nampar anaknyaaa??? doyan pentokin kepala anak ke kaca?? suka mukulin anaknyaaa ampe ungu??? suka nyambit pake <i>belt</i> kulit???”	<i>Physical Threats</i>
2.	@nur*****	“bye muka burik dan bopeng ituuu ih. belajar ye wahai anjing dan tolol biar bisa dpt kerja yg mapan.”	<i>Purposeful Embarrassment</i>
3.	@suk*****	“Man sya mau tanya ada kah bukti yg lo sebut kadrun itu memperbudak WNI? Kalau Cina sipit banyak buktinya.”	<i>Racist</i>
4.	@dav*****	“Si ibu sudah lama gak dipake kali ya, makanya kayak anjing. Keroyok aja bagusnyaaa, perkosa rame rame biar kapok!!!”	<i>Sexual Harassment</i>

3.4 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini adalah variabel independen (x) berupa prediktor atau fitur (kata) terkait *Cyber Harassment* dengan jumlah 5.728 kata atau variabel, di mana setiap kata akan dihitung frekuensi kemunculannya sehingga memiliki nilai 1, 2, 3, dan seterusnya sesuai jumlah kemunculan kata tersebut. Sedangkan variabel dependen (y) berupa empat kelas jenis-jenis *Cyber Harassment*, di mana untuk kelas *Physical Threats* direpresentasikan dengan angka nol (0), *Purposeful Embarrassment* direpresentasikan dengan angka satu (1), *Racist* direpresentasikan dengan angka (2), dan *Sexual Harassment* direpresentasikan dengan angka (3).

Tabel 3.4 Variabel *Cyber Harassment*

No.	Variabel		Jenis	Keterangan
1.	x	<i>Tweet Cyber Harassment</i>	Teks	-
2.	y_1	<i>Physical Threats</i>	Kategori	0
	y_2	<i>Purposeful Embarrassment</i>	Kategori	1
	y_3	<i>Racist</i>	Kategori	2
	y_4	<i>Sexual Harassment</i>	Kategori	3

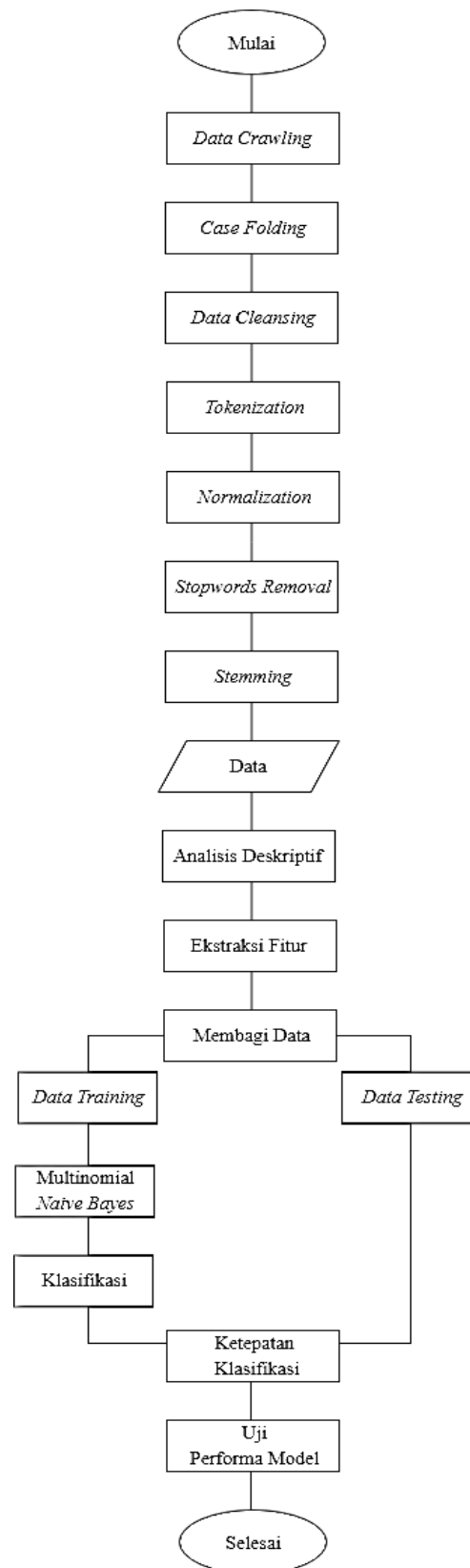
3.5 Tahapan Penelitian

Tahapan penelitian yang dilakukan dalam mengklasifikasi jenis *Cyber Harassment* di media sosial Twitter dengan menggunakan metode Multinomial *Naïve Bayes* adalah sebagai berikut:

1. Mengumpulkan data dengan melakukan *crawling* di media sosial Twitter, kemudian melakukan pelabelan manual pada data ke dalam empat jenis *Cyber Harassment* yaitu *Physical Threats*, *Purposeful Embarrassment*, *Racist*, dan *Sexual Harassment*.
2. Melakukan *data preprocessing* dengan enam tahapan sebagai berikut:
 - a) *Case Folding* untuk menyeragamkan huruf menjadi huruf kecil;
 - b) *Cleansing* untuk membersihkan data dari atribut yang tidak diperlukan;
 - c) *Tokenization* untuk mengubah kalimat menjadi kata-kata yang terpisah;
 - d) *Normalization* untuk menyeragamkan kata yang memiliki makna yang sama;
 - e) *Stopwords removal* untuk menghapus kata hubung, imbuhan, dan sebagainya; dan
 - f) *Stemming* untuk mengubah setiap kata ke dalam bentuk kata dasar.
3. Menganalisis data secara deskriptif untuk mengetahui gambaran umum mengenai kasus *Cyber Harassment* di media sosial Twitter pada tahun 2023.

4. Ekstraksi fitur sebagai langkah transformasi data teks menjadi data numerik dengan menggunakan *Bag of Words*.
5. Membagi *dataset* menjadi data *training* dan data *testing*.
6. Pembuatan model klasifikasi dengan algoritma *Multinomial Naïve Bayes*:
 - a) Menghitung probabilitas *prior* pada data, yaitu probabilitas setiap kelas menggunakan persamaan (2.12);
 - b) Menghitung probabilitas *likelihood* pada data, yaitu probabilitas setiap fitur (kata) pada masing-masing kelas menggunakan persamaan (2.22);
 - c) Menghitung probabilitas *posterior* dengan mengalikan probabilitas *prior* dengan *likelihood* menggunakan persamaan (2.11);
 - d) Mencari nilai probabilitas *posterior* tertinggi untuk menentukan kelas yang terpilih.
7. Menghitung ketepatan hasil klasifikasi dengan menghitung hasil klasifikasi yang tepat dengan *accuracy* menggunakan persamaan (2.23) dan hasil klasifikasi yang tidak tepat dengan *Apparent Error Rate* (APER) menggunakan persamaan (2.24).
8. Melakukan uji performa model dengan *K-fold Cross Validation*.

3.6 Diagram Alir



Gambar 3.1 Diagram Alir Penelitian

BAB IV

HASIL DAN PEMBAHASAN

4.1 *Data Crawling*

Pengumpulan data pada penelitian ini menggunakan teknik *data crawling*, di mana *data crawling* akan mengunduh *tweet* secara otomatis sesuai dengan kata kunci yang telah ditentukan. Pada Twitter, *data crawling* dapat dilakukan apabila telah terdaftar pada Twitter *Developer* dan mendapatkan kode atau *key* sebagai syarat untuk mengunduh atau mengumpulkan data yang bersumber dari Twitter. Penelitian ini mengumpulkan data *tweet* terkait jenis-jenis *Cyber Harassment* dengan menggunakan kata kunci yang sudah ditentukan pada tabel 3.1. Pengumpulan data dengan *data crawling* dilakukan dengan bantuan pemrograman *R*, *script data crawling* terlampir pada lampiran (5).

Pada penelitian ini, *data crawling* dilakukan secara berulang-ulang hingga mendapatkan 2.000 data yang dapat dikategorikan sebagai empat jenis perilaku *Cyber Harassment*, yaitu *Physical Threats*, *Purposeful Embarrassment*, *Racist*, dan *Sexual Harassment*. Pengulangan dalam proses mengumpulkan data dengan *data crawling* dilakukan karena terbatasnya unduhan data yang dapat dilakukan pada media sosial Twitter. Selain itu, kata kunci yang digunakan untuk mengumpulkan data pada penelitian ini akan menarik atau mengunduh seluruh *tweet* yang mengandung kata kunci tersebut, termasuk *tweet* berita, iklan, hingga opini yang bersifat netral. Sehingga, terdapat banyak *tweet* yang terunduh tetapi bukan termasuk ke dalam perilaku *Cyber Harassment*. Oleh karena itu, perlu dilakukan *filtering* terlebih dahulu untuk memberikan label atau kelas secara manual pada data yang telah didapatkan, di mana proses *filtering* tersebut bertujuan menyaring *tweet*

yang tidak berkaitan dengan jenis-jenis *Cyber Harassment*. Jika *tweet* tersebut merupakan perilaku *Cyber Harassment*, maka akan diberikan label atau kelas sesuai empat jenis *Cyber Harassment* yaitu *Physical Threats*, *Purposeful Embarrassment*, *Racist*, dan *Sexual Harassment*, sedangkan apabila *tweet* tersebut bukan merupakan perilaku *Cyber Harassment* maka akan diberikan label “netral”. Tabel 4.1 menunjukkan hasil data *crawling* dan pemberian kelas terhadap setiap data yang terunduh.

Tabel 4.1 Hasil *Data Crawling* dan Pemberian Kelas

No.	<i>Tweet</i>	Kelas
1.	“@nandyradjaanrga @HafidHajizi @guusprakoso Bacot nan, gw Jambak loe”	<i>Physical Threats</i>
2.	"Barangsiapa berwudhu dengan baik keluarlah dosa-dosanya dari jasadnya sampaipun dari bawah kuku-kukunya. (HR. Muslim) #LazismuPiyungan”	Netral
3.	“@OpaJoe_de_Queen @RomeoWalker19 Dia ngomong Anies, kok kedengarannya anjing sich... 🐶”	<i>Purposeful Embarrassment</i>
4.	“@BuniYani Lincoln memelihara jenggot tp ga disebut Radikal, intoleran ato bahkan terorist krn dlm otaknya ga ada ajaran dan pemikiran rasis serta biadab kyk islam! Paham kadrun dungu!??”	<i>Racist</i>
5.	“@sangyoeni [+257] jangan ge, saya orang cina gak paham ⇨ ⇨ ⇨”	Netral
6.	“@nctzenbase Mukenah(muslim),Powerbank, obat darurat, tissue, bawa permen/cemilan manis krna pasti pusing diantara ribuan orang gitu.”	Netral
7.	“@ArnoldPoernomo Its okay @Freya_JKT48 jangan dengarkan chef minoritas ini. Dasar tenaga kerja cina @ArnoldPoernomo”	<i>Racist</i>
8.	“@Lintas_Jaman Gimana jk dia yg kita perkosa mudah ² an dia ikhlas dan mungkin minta nambah 😊😊😏”	<i>Sexual Harassment</i>
9.	"Hasil Pengeluaran HAWAII Hari Ini : Jumat 27-01-2023 Result : 1 2 6 6 SHIO : ANJING Salam JP Kepada pemenang"	Netral
10.	“@cobeh2022 Otak tolol kalian gak nyampe dgn apa maksud beliau.barisan kaum pengikut dajjal kayak kalian tuh tolol”	<i>Purposeful Embarrassment</i>

4.2 *Data Preprocessing*

Data preprocessing adalah langkah penting yang harus dilakukan sebelum mengklasifikasi data. *Data preprocessing* merupakan proses untuk membersihkan data dari *noise*, duplikasi, hingga menghilangkan atribut-atribut yang tidak diperlukan agar dapat memperjelas fitur dan memudahkan model untuk mengklasifikasikan data pada kelas yang tepat. *Data preprocessing* memiliki enam tahap, tahap pertama adalah *case folding* untuk menyeragamkan huruf menjadi huruf kecil; tahap kedua yaitu *data cleansing* untuk menghapus atribut-atribut yang tidak dibutuhkan; tahap ketiga yaitu *tokenization* untuk memisahkan kalimat menjadi kata; tahap keempat yaitu *normalization* untuk menyeragamkan kata yang memiliki arti yang sama; tahap kelima yaitu *stopwords removal* untuk menghapus kata hubung, imbuhan, dan lain sebagainya yang tidak diperlukan; dan tahap terakhir yaitu *stemming* untuk mengubah setiap kata ke dalam bentuk kata dasarnya. *Data preprocessing* pada penelitian ini menggunakan bantuan pemrograman *Python*, *script data preprocessing* terlampir pada lampiran (6).

Pada tabel 4.2 menunjukkan perbedaan data penelitian sebelum dan sesudah melalui *data preprocessing*, selengkapnya hasil *data preprocessing* terlampir pada lampiran (2). Tabel 4.2 menunjukkan bahwa sebelum *data preprocessing*, data merupakan sebuah kalimat panjang yaitu *tweet* yang diunduh sebagai data penelitian, sedangkan setelah *data preprocessing* data menjadi potongan-potongan kata yang mengikuti kata dasarnya. Sehingga kata-kata tersebut adalah variabel independen yang akan digunakan untuk menentukan kelas (variabel dependen) yang tepat bagi dokumen yang mengandung kata-kata tersebut.

Tabel 4.2 Hasil *Data Preprocessing*

No.	Data Sebelum <i>Preprocessing</i>	Data Setelah <i>Preprocessing</i>
1.	nikah 3 kali dia main tangan??? <i>verbally abusive</i> ?? doyan nampar anaknyaaa?? doyan pentokin kepala anak ke kaca?? suka mukulin anaknya ampe ungu??? suka nyambit pake <i>belt</i> kulit???	nikah, kali, main, tangan, <i>verbally</i> , <i>abusive</i> , doyan, nampar, anak, doyan, pentok, kepala, kaca, suka, mukul, anak, ampe, ungu, suka, nyambit, pakai, <i>belt</i> , kulit
2.	<i>bye</i> muka burik dan bopeng ituuu ih. belajar ye wahai anjing dan tolol biar bisa dpt kerja yg mapan	<i>bye</i> , muka, burik, bopeng, wahai, anjing, tolol, biar, bisa, dapat, kerja, mapan
3.	Man sya mau tanya ada kah bukti yg lo sebut kadrin itu memperbudak WNI? Kalau Cina sipit banyak buktinya	mana, saya, tanya, bukti, lu, sebut, kadrin, budak, wni, kalau, cina, sipit, banyak, bukti
4.	Si ibu sudah lama gak dipake kali ya, makanya kayak anjing. Keroyok aja bagusnya, perkosa rame rame biar kapok!!!	ibu, lama, gak, pake, kali, kayak, anjing, keroyok, bagus, perkosa, rame, biar, kapok
⋮	⋮	⋮
2.000	Dan lu adalah keturunan lonte yg diperkosa anjing buduk	lu, adalah, keturunan, lonte, perkosa, anjing, buduk

4.3 Analisis Deskriptif

Penelitian ini menggunakan data *tweet* terkait *Cyber Harassment* di media sosial Twitter yang berjumlah 2.000 data. Analisis deskriptif dalam penelitian ini bertujuan untuk memberikan gambaran umum mengenai data yang digunakan dalam penelitian. Analisis deskriptif dapat mendeskripsikan hasil penelitian secara menyeluruh, sehingga penelitian dapat lebih mudah dipahami dan dapat ditarik kesimpulan mengenai penelitian yang dilakukan.

4.3.1 Analisis Deskriptif Variabel Independen (x)

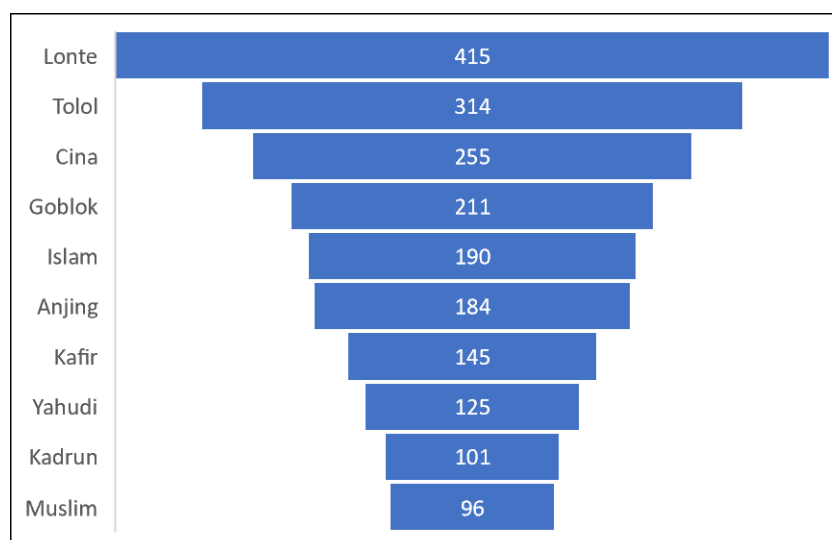
Pada penelitian ini, setiap kata pada data penelitian adalah variabel independen (x) atau variabel prediktor, di mana dari 2.000 data penelitian,

terdapat sebanyak 5.728 kata. Kata pada data penelitian ditunjukkan pada tabel 4.3 sebagai berikut.

Tabel 4.3 Variabel Independen (x)

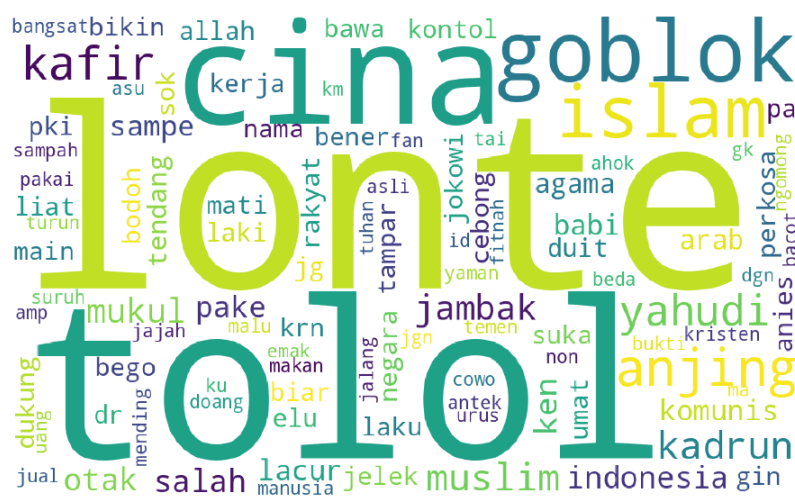
No.	Kata	Variabel
1.	abadi	x_1
2.	abai	x_2
3.	abal	x_3
4.	abang	x_4
5.	abdi	x_5
6.	<i>abuse</i>	x_6
7.	<i>abused</i>	x_7
8.	<i>abusive</i>	x_8
9.	acara	x_9
10.	<i>acting</i>	x_{10}
⋮	⋮	⋮
5.728	zuckerberg	$x_{5.728}$

Pada tabel 4.3 dapat diketahui bahwa seluruh kata yang muncul pada data penelitian akan dihitung nilai probabilitasnya pada setiap kelas, bukan hanya kata yang termasuk kata kunci saja. Dari 5.728 kata yang terdapat pada data, diketahui sepuluh kata dengan frekuensi kemunculan tertinggi sebagaimana yang ditampilkan pada gambar 4.1.



Gambar 4. 1 Frekuensi Kemunculan Kata Tertinggi Pada Kasus *Cyber Harassment*

Pada gambar 4.1, dapat diketahui bahwa sepuluh kata dengan frekuensi kemunculan tertinggi yaitu kata “lonte” muncul 415 kali, kata “tolol” muncul 314 kali, kata “cina” muncul 255 kali, kata “goblok” muncul 211 kali, kata “Islam” muncul 190 kali, kata “anjing” muncul 184 kali, kata “kafir” muncul 145 kali, kata “yahudi” muncul 125 kali, dan kata “muslim” muncul 96 kali, merupakan kata yang mewakili kata kunci yang digunakan dalam mengumpulkan data penelitian. Hal ini dikarenakan pada setiap *tweet* atau dokumen setidaknya terdapat satu kata kunci yang merepresentasikan perilaku *Cyber Harassment*. Sedangkan satu kata lainnya yaitu kata “kadrun” yang muncul 101 kali bukan termasuk kata kunci yang digunakan untuk mengumpulkan data penelitian. Hal tersebut mengindikasikan bahwa kata yang mewakili perilaku *Cyber Harassment* tidak selalu sama setiap waktu atau tidak seluruhnya bergantung pada kata kunci yang digunakan, di mana kata “kadrun” juga menjadi kata yang banyak digunakan dalam kasus *Cyber Harassment* terutama pada jenis *Racist*. Pada gambar 4.2 di bawah ini dapat diketahui kata lainnya yang juga muncul pada kasus *Cyber Harassment*.



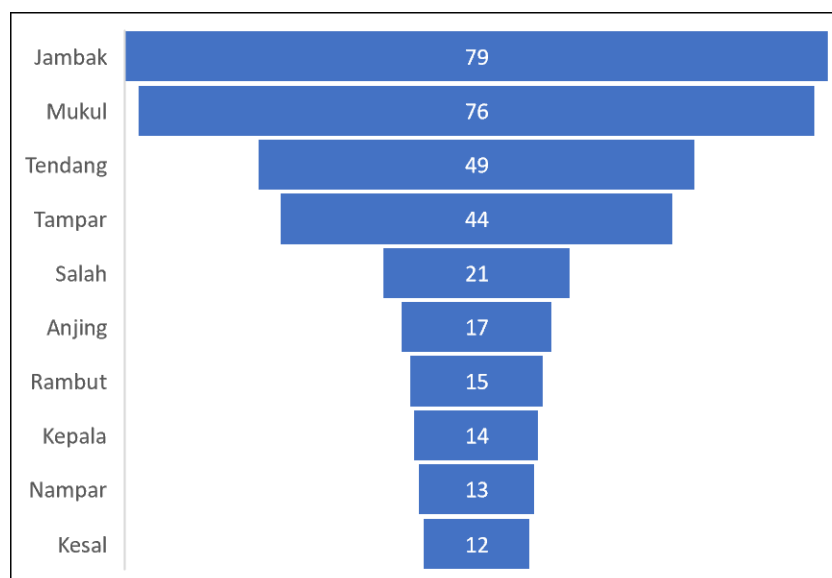
Gambar 4.2 Wordcloud Pada Kasus *Cyber Harassment*

4.3.2 Analisis Deskriptif Variabel Dependen (y)

Pada penelitian ini, empat jenis *Cyber Harassment* yaitu adalah kelas *Physical Threats*, *Purposeful Embarrassment*, *Racist*, dan *Sexual Harassment* adalah variabel dependen (y) atau variabel target.

1. Variabel *Physical Threats* (y_1)

Variabel *Physical Threats* (y_1) merupakan kelas pertama dari jenis *Cyber Harassment*. Setelah melalui *data preprocessing*, diketahui sepuluh kata dengan frekuensi kemunculan tertinggi pada kelas *Physical Threats* ditunjukkan pada gambar 4.3.

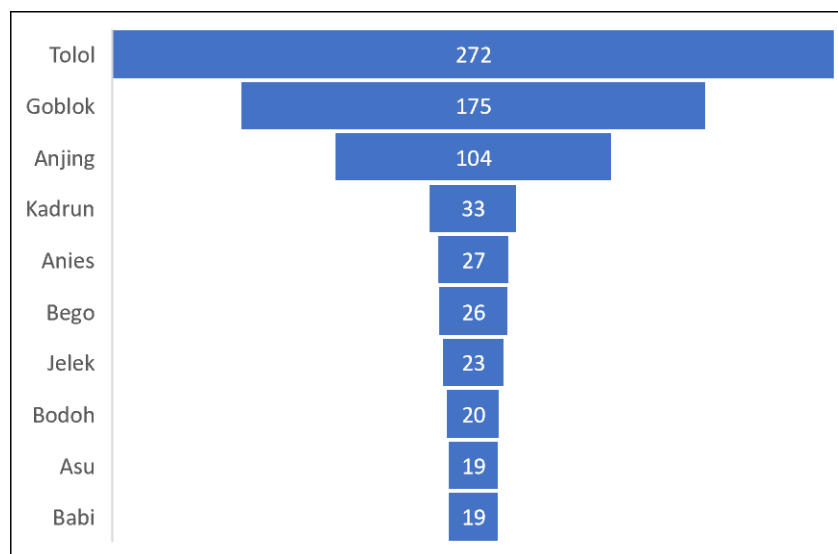


Gambar 4.3 Frekuensi Kemunculan Kata Tertinggi pada Kelas *Physical Threats*

Pada gambar 4.3 menunjukkan sepuluh kata dengan frekuensi tertinggi pada kelas *Physical Threats*, di mana lima kata di antaranya yaitu “jambak” dengan frekuensi kemunculan 79 kali, “mukul” dengan frekuensi kemunculan 76 kali, “tendang” dengan frekuensi kemunculan 49 kali, “tampar” dengan frekuensi kemunculan 44 kali, dan “nampar” dengan frekuensi kemunculan 13 kali, merupakan kata kunci yang digunakan dalam

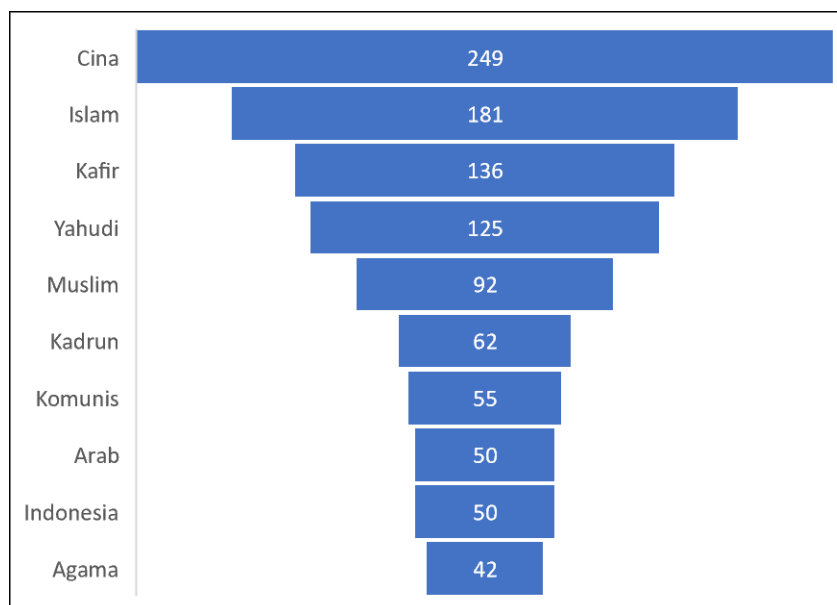
2. Variabel *Purposeful Embarrassment* (y_2)

Variabel *Purposeful Embarrassment* (y_2) merupakan kelas kedua dari jenis *Cyber Harassment*. Setelah melalui *data preprocessing*, diketahui sepuluh kata dengan frekuensi kemunculan tertinggi pada kelas *Purposeful Embarrassment* ditunjukkan pada gambar 4.5



Gambar 4.5 Frekuensi Kemunculan Kata Tertinggi pada Kelas *Purposeful Embarrassment*

Pada gambar 4.5 menunjukkan sepuluh kata dengan frekuensi tertinggi pada kelas *Purposeful Embarrassment*, di mana tujuh kata di antaranya yaitu “tolol” dengan frekuensi kemunculan 272 kali, “goblok” dengan frekuensi kemunculan 175 kali, “anjing” dengan frekuensi kemunculan 104 kali, “bego” dengan frekuensi kemunculan 26 kali, “jelek” dengan frekuensi kemunculan 23 kali, “bodoh” dengan frekuensi kemunculan 20 kali, dan “asu” dengan frekuensi kemunculan 19 kali, merupakan kata kunci yang digunakan dalam mengumpulkan data untuk kelas *Purposeful Embarrassment*. Sedangkan tiga kata lainnya bukan merupakan kata kunci, di mana dua di antaranya yaitu kata “kadrun” dengan frekuensi kemunculan 33 kali dan “anies” dengan frekuensi kemunculan 27 kali merupakan kata



Gambar 4.7 Frekuensi Kemunculan Kata Tertinggi pada Kelas *Racist*

Pada gambar 4.7 menunjukkan sepuluh kata dengan frekuensi tertinggi pada kelas *Racist*, di mana lima kata di antaranya yaitu “cina” dengan frekuensi kemunculan 249 kali, “Islam” dengan frekuensi kemunculan 181 kali, “kafir” dengan frekuensi kemunculan 136 kali, “yahudi” dengan frekuensi kemunculan 125 kali, dan “muslim” dengan frekuensi kemunculan 92 kali, merupakan kata kunci yang digunakan dalam mengumpulkan data untuk kelas *Racist*. Sedangkan lima kata lainnya bukan merupakan kata kunci, di mana tiga kata di antaranya yaitu “kadrin” dengan frekuensi kemunculan 62 kali, “komunis” dengan frekuensi kemunculan 55 kali, dan “arab” dengan frekuensi kemunculan 50 kali, merupakan kata yang sering digunakan pada tindakan rasisme yang menyerang pada perbedaan ideologi dan suku. Hal ini menunjukkan bahwa kata yang mewakili perilaku *Racist* di ruang siber tidak selalu sama setiap waktu atau tidak seluruhnya bergantung pada kata kunci yang digunakan. Bahkan, terdapat kata lain yang mengindikasikan perilaku tercela lainnya seperti peperangan politik

Pada gambar 4.9 menunjukkan sepuluh kata dengan frekuensi tertinggi pada kelas *Sexual Harassment*, di mana empat kata di antaranya yaitu “lonte” dengan frekuensi kemunculan 414 kali, “lacur” dengan frekuensi kemunculan 58 kali, “perkosa” dengan frekuensi kemunculan 46 kali, dan “jalang” dengan frekuensi kemunculan 24 kali, merupakan kata kunci yang digunakan dalam mengumpulkan data untuk kelas *Sexual Harassment*. Sedangkan enam kata lainnya bukan merupakan kata kunci, di mana dua kata di antaranya yaitu “anjing” dengan frekuensi kemunculan 41 kali dan “tolol” dengan frekuensi kemunculan 26 kali merupakan kata kunci pada kelas *Purposeful Embarrassment*, dan dua kata lainnya yaitu “mama” dengan frekuensi kemunculan 24 kali dan “lelaki” dengan frekuensi kemunculan 19 kali dapat merepresentasikan bahwa perilaku *Sexual Harassment* identik dengan menargetkan sosok figur atau jenis kelamin tertentu. Hal ini menunjukkan bahwa kata yang mewakili perilaku *Sexual Harassment* di ruang siber tidak selalu sama setiap waktu atau tidak seluruhnya bergantung pada kata kunci yang digunakan. Bahkan, terdapat kata lain yang mengindikasikan bahwa perilaku *Sexual Harassment* tidak terlepas dari merendahkan dan penghinaan secara verbal. Pada gambar 4.10 di bawah ini, dapat diketahui kata lainnya yang juga muncul pada kelas *Sexual Harassment*.



Gambar 4.10 Wordcloud Pada Kelas *Sexual Harassment*

4.4 Ekstraksi Fitur

Ekstraksi fitur dengan menggunakan model *Bag of Words* (BOW) dilakukan untuk mengubah data teks menjadi data numerik, di mana setiap kata dihitung kemunculannya pada setiap dokumen. BOW tidak mementingkan urutan dan tata bahasa, sehingga nantinya model akan mengklasifikasi kelas yang tepat sesuai angka atau nilai dari kata dalam kalimat untuk setiap dokumen. Tabel 4.4 menunjukkan hasil ekstraksi fitur pada data penelitian, yang selengkapnya terlampir pada lampiran (3).

Tabel 4.4 Hasil Ekstraksi Fitur dengan *Bag of Words*

Dok.	anak	...	bukti	...	cina	...	perkosan	...	rame	...	suka
1.	3	...	0	...	0	...	0	...	0	...	2
2.	0	...	0	...	0	...	0	...	0	...	0
3.	0	...	2	...	1	...	0	...	0	...	0
4.	0	...	0	...	0	...	1	...	2	...	0
5.	0	...	0	...	1	...	0	...	0	...	0
...
2.000	0	...	0	...	0	...	1	...	0	...	0

Tabel 4.4 menunjukkan frekuensi kata dari setiap kata pada setiap dokumen, di mana jika kata tidak ada di dalam sebuah dokumen akan muncul nilai 0,

sedangkan jika muncul maka akan diberi nilai 1, 2, 3, dan seterusnya sesuai jumlah kemunculannya.

4.5 Analisis Metode Multinomial *Naïve Bayes*

4.5.1 Pembagian Data *Training* dan Data *Testing*

Dalam melakukan klasifikasi dengan metode Multinomial *Naïve Bayes*, perlu dilakukannya pembagian data menjadi *data training* dan *data testing*. Pembagian data ini bertujuan untuk mengetahui nilai akurasi terbaik, di mana pada umumnya, semakin banyak *data training* sebagai data pembelajaran bagi model, maka akan semakin baik nilai akurasi dari hasil klasifikasi model pada *data testing*. Pembagian data pada penelitian ini dibagi ke dalam enam kombinasi seperti pada tabel 4.5.

Tabel 4.5 Pembagian *Data Training* dan *Data Testing*

No.	Persentase Data (%)		Akurasi (%)	
	<i>Data Training</i>	<i>Data Testing</i>	Tabulasi Silang	<i>K-fold Cross Validation</i>
1.	70	30	73	74,58
2.	75	25	75	75,07
3.	80	20	74	75,55
4.	85	15	75	75,67
5.	90	10	77	75,78
6.	95	5	77	76,21

Berdasarkan tabel 4.5 didapatkan bahwa nilai akurasi dari ke-enam kombinasi dikategorikan sebagai nilai akurasi yang baik pada angka 74,58% hingga 76,21%. Kombinasi *data training* 95% dan *data testing* 5% memiliki nilai akurasi yang paling baik dibandingkan dengan kombinasi lainnya, yaitu sebesar 76,21%. Oleh karena itu, penelitian ini menggunakan *data training* 95% sejumlah 1.900 data dan *data testing* 5% sejumlah 100 data.

4.5.2 Klasifikasi Jenis-Jenis *Cyber Harassment* Menggunakan Metode *Multinomial Naïve Bayes*

Setelah menentukan kombinasi pembagian *data training* dan *data testing*, selanjutnya data yang telah dibagi tersebut dapat digunakan untuk mengklasifikasi dengan metode *Multinomial Naïve Bayes*. Pada tabel 4.6 ditampilkan 5 dokumen untuk digunakan sebagai contoh dalam proses klasifikasi dengan model *Multinomial Naïve Bayes*, di mana model akan mengklasifikasi kelas yang tepat untuk dokumen ke-5 pada *data testing*.

Tabel 4.6 Contoh *Dataset* untuk Klasifikasi

	Dok.	Kata	Kelas	Hasil Klasifikasi
<i>Data Training</i>	1.	nikah kali main tangan <i>verbally abusive</i> doyan nampar anak doyan pentok kepala kaca suka mukul anak ampe ungu suka nyambit pakai <i>belt</i> kulit	<i>Physical Threats</i>	<i>Physical Threats</i>
	2.	<i>bye</i> muka burik bopeng wahai anjing tolol biar bisa dapat kerja mapan	<i>Purposeful Embarrassment</i>	<i>Purposeful Embarrassment</i>
	3.	mana saya tanya bukti lu sebut kadrin budak wni kalau cina sipit banyak bukti	<i>Racist</i>	<i>Racist</i>
	4.	ibu lama gak pake kali kayak anjing keroyok bagus perkosa rame biar kapok	<i>Sexual Harassment</i>	<i>Sexual Harassment</i>
<i>Data Testing</i>	5.	lu adalah keturunan lonte perkosa anjing buduk	<i>Sexual Harassment</i>	?

Terdapat empat langkah untuk mengklasifikasi dengan Multinomial *Naïve Bayes* seperti yang tertera pada tahap penelitian sub-bab 3.5 pada poin nomor 6, yaitu sebagai berikut:

1. Menentukan Probabilitas *Prior* Kelas *Cyber Harassment*

Berdasarkan tabel 4.6, diketahui bahwa dari 5 dokumen terdapat 1 dokumen dengan kelas *Physical Threats*, 1 dokumen dengan kelas *Purposeful Embarrassment*, 1 dokumen dengan kelas *Racist*, dan 2 dokumen dengan kelas *Sexual Harassment*. Untuk menghitung probabilitas *prior* untuk setiap kelas dapat menggunakan persamaan (2.12). Tabel 4.7 menunjukkan nilai probabilitas *prior* untuk setiap kelas pada jenis-jenis *Cyber Harassment*.

Tabel 4.7 Probabilitas *Prior* Contoh *Dataset* untuk Klasifikasi

Probabilitas			
<i>Physical Threats</i>	<i>Purposeful Embarrassment</i>	<i>Racist</i>	<i>Sexual Harassment</i>
$\frac{1}{5} = 0,2$	$\frac{1}{5} = 0,2$	$\frac{1}{5} = 0,2$	$\frac{2}{5} = 0,4$

2. Menentukan Nilai *Likelihood* Setiap Kata Pada Setiap Kelas *Cyber Harassment*

Setelah menghitung nilai probabilitas *prior* dari setiap kelas, selanjutnya akan dihitung nilai *likelihood* untuk variabel independen (x) atau setiap kata pada dokumen pada masing-masing kelas (y) menggunakan persamaan (2.22). Berdasarkan tabel 4.6, maka didapatkan informasi sebagaimana ditunjukkan pada tabel 4.8 sebagai berikut.

Tabel 4.8 Contoh *Dataset* untuk Klasifikasi

	Dok.	Kata	Jumlah Kata (N)
<i>Data Training</i>	1.	nikah kali main tangan <i>verbally abusive</i> doyan nampar anak doyan pentok kepala kaca suka mukul anak ampe ungu suka nyambit pakai <i>belt</i> kulit	23
	2.	<i>bye</i> muka burik bopeng wahai anjing tolol biar bisa dapat kerja mapan	12
	3.	mana saya tanya bukti lu sebut kadrin budak wni kalau cina sipit banyak bukti	14
	4.	ibu lama gak pake kali kayak anjing keroyok bagus perkosa rame biar kapok	13
<i>Data Testing</i>	5.	lu adalah keturunan lonte perkosa anjing buduk	7

Pada tabel 4.8 diketahui jumlah kata yang muncul pada setiap dokumen (N). Jumlah kata yang muncul pada dokumen 1 adalah 23 kata, pada dokumen 2 adalah 12 kata, pada dokumen 3 adalah 14 kata, pada dokumen 4 adalah 13 kata, dan pada dokumen 5 adalah 7 kata. Berdasarkan informasi tersebut, maka dapat diketahui jumlah kata berbeda atau *vocabulary* (α_d) pada seluruh dokumen pada tabel 4.9.

Tabel 4.9 Jumlah Kata Berbeda (*Vocabulary*)

Daftar Kata						Jumlah Kata Berbeda (α_d)
<i>abusive</i>	adalah	ampe	anak	anjing	bagus	59
banyak	<i>belt</i>	biar	bisa	bopeng	budak	
buduk	bukti	burik	<i>bye</i>	cina	dapat	
doyan	gak	ibu	kaca	kadrin	kalau	
kali	kapok	kayak	kepala	kerja	keroyok	
keturunan	kulit	lama	lonte	lu	main	

Daftar Kata						Jumlah Kata Berbeda (α_d)
mana	mapan	muka	mukul	nampar	nikah	59
nyambit	pakai	pake	pentok	perkosa	rame	
saya	sebut	sipit	suka	tangan	tanya	
tolol	ungu	<i>verbally</i>	wahai	wni	tolol	

Tabel 4.9 merupakan daftar kata atau *vocabulary* yaitu kemunculan kata yang berbeda pada seluruh dokumen yaitu dokumen 1-5 yang berjumlah 59 kata. Setelah mengetahui jumlah kemunculan kata pada setiap dokumen (N) dan jumlah seluruh kata berbeda pada seluruh dokumen (α_d), maka dapat dicari nilai ekstraksi fitur dari dokumen ke-5. Pada tabel 4.10 di bawah menunjukkan ekstraksi fitur dari dokumen ke-5 pada pada tabel 4.6.

Tabel 4.10 Ekstraksi Fitur Dokumen 5

Kata Pada Dokumen 5							
Dok.	lu	adalah	keturunan	lonte	perkosa	anjing	buduk
1.	0	0	0	0	0	0	0
2.	0	0	0	0	0	1	0
3.	1	0	0	0	0	0	0
4.	0	0	0	0	1	1	0
5.	0	0	0	0	0	0	0

Tabel 4.10 menunjukkan bahwa kata “lu” muncul 1 kali pada dokumen ke-3, kata “perkosa” muncul 1 kali pada dokumen ke-4, kata “anjing” muncul 1 kali pada dokumen ke-2 dan ke-4. Berdasarkan pemaparan tersebut, maka dengan menggunakan persamaan (2.22) dapat dihitung nilai *likelihood* untuk setiap kata pada dokumen 5 terhadap masing-masing kelas. Tabel 4.11 merupakan tabel perhitungan untuk nilai *likelihood* setiap kata pada setiap kelas.

Tabel 4.11 Perhitungan Nilai *Likelihood* $P(W_i|C)$

<i>Physical Threats</i>	<i>Purposeful Embarrassment</i>
$P(lu PT)$	$P(lu PE)$
$= \frac{0 + 1}{23 + 59} = 0,0121$	$= \frac{0 + 1}{12 + 59} = 0,0140$
$P(adalah PT)$	$P(adalah PE)$
$= \frac{0 + 1}{23 + 59} = 0,0121$	$= \frac{0 + 1}{12 + 59} = 0,0140$
$P(keturunan PT)$	$P(keturunan PE)$
$= \frac{0 + 1}{23 + 59} = 0,0121$	$= \frac{0 + 1}{12 + 59} = 0,0140$
$P(lonte PT)$	$P(lonte PE)$
$= \frac{0 + 1}{23 + 59} = 0,0121$	$= \frac{0 + 1}{12 + 59} = 0,0140$
$P(perkosa PT)$	$P(perkosa PE)$
$= \frac{0 + 1}{23 + 59} = 0,0121$	$= \frac{0 + 1}{12 + 59} = 0,0140$
$P(anjing PT)$	$P(anjing PE)$
$= \frac{0 + 1}{23 + 59} = 0,0121$	$= \frac{1 + 1}{12 + 59} = 0,0281$
$P(buduk PT)$	$P(buduk PE)$
$= \frac{0 + 1}{23 + 59} = 0,0121$	$= \frac{0 + 1}{12 + 59} = 0,0140$
<i>Racist</i>	<i>Sexual Harassment</i>
$P(lu R)$	$P(lu SH)$
$= \frac{1 + 1}{14 + 59} = 0,0273$	$= \frac{0 + 1}{13 + 59} = 0,0138$
$P(adalah R)$	$P(adalah SH)$
$= \frac{0 + 1}{14 + 59} = 0,0136$	$= \frac{0 + 1}{13 + 59} = 0,0138$
$P(keturunan R)$	$P(keturunan SH)$
$= \frac{0 + 1}{14 + 59} = 0,0136$	$= \frac{0 + 1}{13 + 59} = 0,0138$
$P(lonte R)$	$P(lonte SH)$
$= \frac{0 + 1}{14 + 59} = 0,0136$	$= \frac{0 + 1}{13 + 59} = 0,0138$
$P(perkosa R)$	$P(perkosa SH)$
$= \frac{0 + 1}{14 + 59} = 0,0136$	$= \frac{1 + 1}{13 + 59} = 0,0277$
$P(anjing R)$	$P(anjing SH)$
$= \frac{0 + 1}{14 + 59} = 0,0136$	$= \frac{1 + 1}{13 + 59} = 0,0277$
$P(buduk R)$	$P(buduk SH)$
$= \frac{0 + 1}{14 + 59} = 0,0136$	$= \frac{0 + 1}{13 + 59} = 0,0138$

3. Menentukan Probabilitas Bersyarat Setiap Kelas *Cyber Harassment*

Langkah selanjutnya adalah menghitung probabilitas bersyarat atau *posterior* untuk setiap variabel independen (x) terhadap setiap kelas (y) dengan mengalikan nilai *prior* dan *likelihood* pada masing-masing kelas menggunakan persamaan (2.11) sebagai berikut:

1. Kelas *Physical Threats*

$$\begin{aligned}
 & P(\text{Physical Threats} | \text{Dok. 5}) \\
 &= P(C = c_1) (P(\text{lu} | PT) \times P(\text{adalah} | PT) \times P(\text{keturunan} | PT) \times \\
 & P(\text{lonte} | PT) \times P(\text{perkosa} | PT) \times P(\text{anjing} | PT) \times P(\text{buduk} | PT)) \\
 &= (0,2) ((0,0121) \times (0,0121) \times (0,0121) \times (0,0121) \times \\
 & (0,0121) \times (0,0121) \times (0,0121)) \\
 &= 7,5949 \times 10^{-15}
 \end{aligned}$$

2. Kelas *Purposeful Embarrassment*

$$\begin{aligned}
 & P(\text{Purposeful Embarrassment} | \text{Dok. 5}) \\
 &= P(C = c_2) (P(\text{lu} | PE) \times P(\text{adalah} | PE) \times P(\text{keturunan} | PE) \times \\
 & P(\text{lonte} | PE) \times P(\text{perkosa} | PE) \times P(\text{anjing} | PE) \times P(\text{buduk} | PE)) \\
 &= (0,2) ((0,0140) \times (0,0140) \times (0,0140) \times (0,0140) \times \\
 & (0,0140) \times (0,0281) \times (0,0140)) \\
 &= 4,3787 \times 10^{-14}
 \end{aligned}$$

3. Kelas *Racist*

$$\begin{aligned}
 & P(\text{Racist} | \text{Dok. 5}) \\
 &= P(C = c_3) (P(\text{lu} | R) \times P(\text{adalah} | R) \times P(\text{keturunan} | R) \times \\
 & P(\text{lonte} | R) \times P(\text{perkosa} | R) \times P(\text{anjing} | R) \times P(\text{buduk} | R))
 \end{aligned}$$

$$\begin{aligned}
&= (0,2)((0,0273) \times (0,0136) \times (0,0136) \times (0,0136) \times \\
&(0,0136) \times (0,0136) \times (0,0136)) \\
&= 3,4548 \times 10^{-14}
\end{aligned}$$

4. Kelas *Sexual Harassment*

$$\begin{aligned}
&P(\text{Sexual Harassment} | \text{Dok. 5}) \\
&= P(C = c_4)(P(\text{lu} | SH) \times P(\text{adalah} | SH) \times P(\text{keturunan} | SH) \times \\
&P(\text{lonte} | SH) \times P(\text{perkosa} | SH) \times P(\text{anjing} | SH) \times P(\text{buduk} | SH)) \\
&= (0,4)((0,0138) \times (0,0138) \times (0,0138) \times (0,0138) \times \\
&(0,0277) \times (0,0277) \times (0,0138)) \\
&= 1,5360 \times 10^{-11}
\end{aligned}$$

4. Menentukan Kelas dengan Nilai Probabilitas Bersyarat Tertinggi

Langkah untuk menentukan kelas yang tepat untuk dokumen ke-5 yaitu dengan menentukan kelas dengan nilai probabilitas bersyarat atau *posterior* tertinggi. Hasil perhitungan nilai probabilitas bersyarat atau *posterior* ditunjukkan pada tabel 4.12.

Tabel 4.12 Probabilitas Bersyarat

Probabilitas Bersyarat			
<i>Physical Threats</i>	<i>Purposeful Embarrassment</i>	<i>Racist</i>	<i>Sexual Harassment</i>
$7,5949 \times 10^{-15}$	$4,3787 \times 10^{-14}$	$3,4548 \times 10^{-14}$	$1,5360 \times 10^{-11}$

Berdasarkan tabel 4.12 diketahui bahwa nilai probabilitas bersyarat untuk kelas *Sexual Harassment* tepat untuk dokumen 5 adalah $1,5360 \times 10^{-11}$, nilai probabilitas kelas *Purposeful Embarrassment* tepat untuk dokumen 5 adalah $4,3787 \times 10^{-14}$, nilai probabilitas kelas *Racist* tepat

untuk dokumen 5 adalah $3,4548 \times 10^{-14}$, dan nilai probabilitas kelas *Physical Threats* tepat untuk dokumen 5 adalah $7,5949 \times 10^{-15}$. Sehingga dapat disimpulkan bahwa dokumen ke-5 terklasifikasi sebagai kelas *Sexual Harassment*.

4.5.3 Hasil Klasifikasi Jenis-Jenis *Cyber Harassment* Menggunakan Metode Multinomial *Naïve Bayes*

Multinomial *Naïve Bayes* akan menentukan kelas yang tepat untuk setiap dokumen sebagaimana langkah-langkah yang telah dijelaskan pada poin sebelumnya. Pada penelitian ini, klasifikasi jenis-jenis *Cyber Harassment* di media sosial Twitter dengan metode Multinomial *Naïve Bayes* menggunakan bantuan pemrograman *Python*, *script* Multinomial *Naïve Bayes* terlampir pada lampiran (7). Tabel 4.13 merupakan hasil klasifikasi kelas yang tepat untuk jenis-jenis *Cyber Harassment* di media sosial Twitter, selengkapnya terlampir pada lampiran (4).

Tabel 4.13 Hasil Klasifikasi *Data Testing* Menggunakan Multinomial *Naïve Bayes*

Dok.	Kata	Kelas
1.	panji gumilang memang anak anjing pki biadab panji gumilang	<i>Racist</i>
2.	otak tolol kalian gak nyampe dgn maksud beliau baris kaum ikut dajjal kayak kalian tolol	<i>Purposeful Embarrassment</i>
3.	lonte baru nonton bola diem	<i>Sexual Harassment</i>
4.	uang pusat bebas pemda dki blok goblok itu wilayah dki lahan diskusi kpda rakyat tinggal bayar tolol km blok	<i>Purposeful Embarrassment</i>
5.	pukul pake gantung jemur el	<i>Physical Threats</i>
⋮	⋮	⋮
99.	cina parasit parasit ekonomi	<i>Racist</i>
100.	ngomong anies kedengar anjing sich	<i>Purposeful Embarrassment</i>

Jumlah hasil klasifikasi setiap kelas yang diperoleh model dengan metode Multinomial *Naïve Bayes* adalah sebagai berikut:

Tabel 4.14 Jumlah Hasil Klasifikasi *Cyber Harassment* Menggunakan Multinomial *Naïve Bayes* Pada Setiap Kelas

<i>Cyber Harassment</i>			
<i>Physical Threats</i>	<i>Purposeful Embarrassment</i>	<i>Racist</i>	<i>Sexual Harassment</i>
27	19	31	23

Tabel 4.14 menunjukkan bahwa dari 5% atau 100 *data testing* yang digunakan untuk mengklasifikasi data *tweet Cyber Harassment* di media sosial Twitter, terdapat 27 dokumen yang terklasifikasi sebagai kelas *Physical Threats*, 19 dokumen sebagai kelas *Purposeful Embarrassment*, 31 dokumen sebagai kelas *Racist*, dan 23 dokumen sebagai kelas *Sexual Harassment*. Hasil tersebut menunjukkan bahwa kelas dengan hasil klasifikasi paling banyak adalah kelas *Racist* dengan jumlah 31 dokumen dan kelas dengan hasil klasifikasi paling sedikit adalah kelas *Purposeful Embarrassment* dengan jumlah 19 dokumen.

4.5.4 Ketepatan Klasifikasi

Ketepatan klasifikasi dapat diketahui dengan menghitung hasil klasifikasi pada tabulasi silang yang terdiri dari *True Class* dan *False Class*.

Tabel 4.15 merupakan tabel tabulasi silang dengan empat kelas.

Tabel 4.15 Hasil Tabulasi Silang

<i>Class</i>	<i>Predicted class</i>			
	<i>Physical Threats</i>	<i>Purposeful Embarrassment</i>	<i>Racist</i>	<i>Sexual Harassment</i>
<i>Physical Threats</i>	20	4	1	0
<i>Purposeful Embarrassment</i>	5	10	5	0
<i>Racist</i>	1	3	25	1
<i>Sexual Harassment</i>	1	2	0	22

Berdasarkan tabel 4.15 diketahui bahwa nilai *True Class* untuk kelas *Physical Threats* (T_{11}) adalah 20, artinya terdapat sebanyak 20 data *tweet Cyber Harassment* yang diklasifikasi benar sebagai kelas *Physical Threats*; nilai *True Class* untuk kelas *Purposeful Embarrassment* (T_{22}) adalah 10, artinya terdapat sebanyak 10 data *tweet Cyber Harassment* yang diklasifikasi benar sebagai *Purposeful Embarrassment*; nilai *True Class* untuk kelas *Racist* (T_{33}) adalah 25, artinya terdapat sebanyak 25 data *tweet Cyber Harassment* yang diklasifikasi benar sebagai *Racist*; dan nilai *True Class* untuk kelas *Sexual Harassment* (T_{44}) adalah 22, artinya terdapat sebanyak 22 data *tweet Cyber Harassment* yang diklasifikasi benar sebagai *Sexual Harassment*.

Nilai *False Class* untuk kelas *Physical Threats* adalah 4 (F_{12}), 1 (F_{13}), dan 0 (F_{14}), artinya terdapat 4 data *Physical Threats* yang diklasifikasi salah sebagai *Purposeful Embarrassment*, 1 data *tweet Physical Threats* yang diklasifikasi salah sebagai *Racist*, dan 0 data *tweet Physical Threats* yang diklasifikasi salah sebagai *Sexual Harassment*; nilai *False Class* untuk kelas *Purposeful Embarrassment* adalah 5 (F_{21}), 5 (F_{23}), dan 0 (F_{24}), artinya terdapat 5 data *tweet Purposeful Embarrassment* yang diklasifikasi salah sebagai *Physical Threats*, 5 data *tweet Purposeful Embarrassment* yang diklasifikasi salah sebagai *Racist*, dan 0 data *tweet Purposeful Embarrassment* yang diklasifikasi salah sebagai *Sexual Harassment*; nilai *False Class* untuk kelas *Racist* adalah 1 (F_{31}), 3 (F_{32}), dan 1 (F_{34}), artinya terdapat 1 data *tweet Racist* yang diklasifikasi salah sebagai *Physical Threats*, 3 data *tweet Racist* yang diklasifikasi salah sebagai *Purposeful Embarrassment*, dan 1 data *tweet Racist* yang diklasifikasi salah sebagai *Sexual Harassment*; dan nilai *False Class* untuk

kelas *Sexual Harassment* adalah 1 (F_{41}), 2 (F_{42}), dan 0 (F_{43}), artinya terdapat 1 data *tweet Sexual Harassment* yang diklasifikasi salah sebagai *Physical Threats*, 2 data *tweet Sexual Harassment* yang diklasifikasi salah sebagai *Purposeful Embarrassment*, dan 0 data *tweet Sexual Harassment* yang diklasifikasi salah sebagai *Racist*.

Nilai ketepatan klasifikasi dapat dihitung dengan menggunakan persamaan (2.23) sebagai berikut:

$$\begin{aligned} Accuracy &= \frac{20 + 10 + 25 + 22}{20+4+1+0+5+10+5+0+1+3+25+1+1+2+0+22} \times 100\% \\ &= 77\% \end{aligned}$$

Hasil perhitungan tersebut menunjukkan bahwa tingkat ketepatan klasifikasi model Multinomial *Naïve Bayes* dalam mengklasifikasikan 5% atau 100 *data testing* adalah sebesar 77% yang dapat dikategorikan sebagai tingkat ketepatan klasifikasi yang baik. Selain itu, dapat diketahui tingkat *error* atau kesalahan klasifikasi dengan menghitung *APER* (*Apparent Error Rate*) menggunakan persamaan (2.24) sebagai berikut:

$$\begin{aligned} APER &= \frac{4+1+0+5+5+0+1+3+1+1+2+0}{20+4+1+0+5+10+5+0+1+3+25+1+1+2+0+22} \times 100\% \\ &= 23\% \end{aligned}$$

Hasil perhitungan tersebut menunjukkan bahwa tingkat kesalahan klasifikasi model Multinomial *Naïve Bayes* dalam mengklasifikasikan 5% atau 100 *data testing* adalah 23% atau 0,23.

4.6 Uji Performa Model

Dalam menguji performa model pada penelitian ini digunakan metode *K-fold Cross Validation*. Metode ini menjadikan seluruh data yang digunakan dalam penelitian dapat diambil secara acak dan dapat menjadi *data training* maupun *data testing*. Penelitian ini menggunakan $k = 10$ yang artinya data dibagi dalam 10 *fold*.

Tabel 4.16 Hasil *K-fold Cross Validation*

<i>Fold</i> ke-	Akurasi (%)
1	75,09
2	74,82
3	75,62
4	75,94
5	75,94
6	75,62
7	76,21
8	75,89
9	76,16
10	75,94

Tabel 4.16 menunjukkan hasil nilai akurasi pada setiap *fold*, di mana dapat dilihat bahwa akurasi memiliki nilai yang berbeda-beda pada setiap *fold*nya yaitu dengan nilai 75,09% sampai dengan 76,21%. Hal ini dikarenakan *k-fold cross validation* mengambil secara acak data untuk dijadikan *data training* dan *data testing*, apabila *fold* 1 berperan sebagai *data testing*, maka *fold* sisanya $k - 1$ akan berperan sebagai *data training*, dan begitu seterusnya. Berdasarkan tabel 4.16 dapat diketahui bahwa nilai akurasi terbaik berada pada *fold* ke-7 dengan nilai akurasi sebesar 76,21%. Tabel 4.16 juga menunjukkan bahwa hasil dari nilai akurasi adalah konsisten dan menunjukkan nilai yang tidak jauh berbeda pada setiap pengulangan atau *fold*, sehingga dapat diinterpretasikan bahwa model maupun data dalam penelitian dapat dikatakan reliabel.

4.7 Kajian Keislaman dengan Hasil Penelitian

Hasil penelitian menunjukkan bahwa dalam perilaku *Cyber Harassment* yang merupakan perilaku menyimpang dan penyalahgunaan teknologi untuk berbuat kejahatan terhadap sesama manusia, diketahui bahwa setiap kata dapat membentuk kalimat dan melahirkan makna yang bertujuan untuk merendahkan, memermalukan, menyakiti, hingga melecehkan orang lain. Hasil klasifikasi menunjukkan bahwa perilaku *Racist* menjadi perilaku dengan jumlah kasus terbanyak, hal ini menunjukkan betapa manusia sangat menyimpang dari ajaran agama Islam. Tindakan *Racist* sangat bertentangan dengan firman Allah *subhanahu wa ta'ala* dalam surah Ar-Rum ayat 22 sebagai berikut:

“... Dan di antara tanda-tanda kekuasaan-Nya ialah menciptakan langit dan bumi dan berlain-lainan bahasamu dan warna kulitmu. Sesungguhnya pada yang demikian itu benar-benar terdapat tanda-tanda bagi orang-orang yang mengetahui”

Menurut tafsir yang tercantum dalam jurnal Pendidikan Multikultural dalam Perspektif Qur'an oleh Amin dan Erihadiana (2022), surah Ar-Rum ayat 22 memberikan penjelasan bahwa dalam Islam terdapat pendidikan multikultural, di mana sesama manusia harus saling menghargai satu sama lain sebagai makhluk Allah *subhanahu wa ta'ala*, dan tidak ada yang lebih baik di antara yang lainnya berdasarkan perbedaan yang ada. Pada surah Ar-Rum ayat 22 juga ditegaskan bahwa hanya orang-orang yang berpengetahuan yang dapat berpikir dengan benar dan belajar, bahwa perbedaan tersebut adalah *sunnatullah* yang tidak dapat dihindarkan dan harus dihargai karena seluruhnya adalah ciptaan Allah *subhanahu wa ta'ala*.

Islam sebagai agama yang menunjukkan jalan kebenaran menentang segala jenis tindak kejahatan termasuk pelecehan, menyakiti orang lain dengan kekerasan,

hingga memermalukan atau menjatuhkan nama baik seseorang. Allah *subhanahu wa ta'ala* berfirman dalam surah Al-Isra' ayat 70:

"... Dan sungguh, Kami telah memuliakan anak cucu Adam, dan Kami angkut mereka di darat dan di laut, dan Kami beri mereka rezeki dari yang baik-baik dan Kami lebihkan mereka di atas banyak makhluk yang Kami ciptakan dengan kelebihan yang sempurna"

Menurut tafsir Kementerian Agama dalam jurnal Ilmu-ilmu Al-quran Hadits Syari'dah dan Tarbiyah (2021) menjelaskan bahwa maksud dari surah Al-Isra' ayat 70 adalah bahwa Allah *subhanahu wa ta'ala* telah mengilhami manusia sebagai makhluk yang mulia baik itu laki-laki atau perempuan. Allah *subhanahu wa ta'ala* telah memberi rezeki berupa kesehatan, pemahaman, pemikiran, kekuatan, dan sebagainya agar manusia dapat memanfaatkannya dengan baik dan untuk kebaikan, agar manusia saling membantu sama lain. Allah *subhanahu wa ta'ala* menegaskan bahwa baik laki-laki dan perempuan harus saling menjaga satu sama lain dan dilarang untuk menjatuhkan atau menyakiti satu sama lain sebagaimana dalam surah At-Taubah ayat 71.

"Orang-orang mukmin, laki-laki dan perempuan, sebagian mereka menjadi penolong bagi sebagian yang lain. Mereka menyuruh (berbuat) makruf dan mencegah (berbuat) munkar, menegakkan salat, menunaikan zakat, dan taat kepada Allah dan Rasul-Nya. Mereka akan diberi Rahmat oleh Allah. Sesungguhnya Allah Mahaperkasa lagi Mahabijaksana"

Menurut tafsir Kementerian Agama, surah At-Taubah ayat 21 tersebut menegaskan bahwa manusia adalah menjadi penolong bagi manusia lainnya, saling membantu dan meninggalkan perbuatan yang tercela. Sehingga tidak dibenarkan bagi manusia untuk melakukan hal-hal yang tercela seperti kekerasan, memermalukan, mendiskriminasi, hingga melecehkan manusia lainnya, karena itu adalah sifat orang-orang yang tercela. Sebagaimana Allah *subhanahu wa ta'ala* memuliakan ciptaan-Nya yaitu manusia, maka haruslah bagi manusia kepada

sesama manusia lain saling memuliakan dengan berbuat baik dan mengingatkan kepada kebaikan serta mencegah dari keburukan.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan pembahasan mengenai hasil analisis mengenai klasifikasi jenis-jenis *Cyber Harassment* maka diperoleh kesimpulan sebagai berikut:

1. Pendekatan Multinomial *Naïve Bayes* dalam mengklasifikasi jenis-jenis *Cyber Harassment* dengan data dari media sosial Twitter memiliki hasil klasifikasi terbaik pada kombinasi pembagian data 95% atau 1.900 dokumen sebagai *data training* dan 5% atau 100 dokumen sebagai *data testing*. Hasil klasifikasi menggunakan metode Multinomial *Naïve Bayes* dari 100 *data testing* yang digunakan untuk diklasifikasi, diketahui menghasilkan sebanyak 77 data yang diklasifikasi tepat, dengan rincian 20 dokumen diklasifikasi benar sebagai kelas *Physical Threats*, 10 dokumen diklasifikasi benar sebagai kelas *Purposeful Embarrassment*, 25 dokumen diklasifikasi benar sebagai kelas *Racist*, dan 22 dokumen diklasifikasi benar sebagai kelas *Sexual Harassment*. Sedangkan 23 data lainnya diklasifikasi pada kelas yang salah, dengan rincian 7 dokumen diklasifikasi salah sebagai kelas *Physical Threats*, 9 dokumen diklasifikasi salah sebagai kelas *Purposeful Embarrassment*, 6 dokumen diklasifikasi salah sebagai kelas *Racist*, dan 1 dokumen diklasifikasi salah sebagai kelas *Sexual Harassment*.
2. Hasil ketepatan klasifikasi jenis-jenis *Cyber Harassment* dengan metode Multinomial *Naïve Bayes* adalah sebesar 77%, yang artinya model Multinomial *Naïve Bayes* dapat melakukan klasifikasi jenis-jenis *Cyber Harassment* dalam kategori baik. Uji performa model yang dilakukan pada Multinomial *Naïve*

Bayes menggunakan *K-fold Cross Validation* dengan mempartisi data ke dalam 10 *fold* memperoleh nilai akurasi pada presentase 75,09% sampai dengan 76,21%. Hasil uji performa model tersebut menunjukkan bahwa pada *fold* ke-7, didapatkan partisi data terbaik hingga menghasilkan tingkat akurasi sebesar 76,21%. Sehingga dapat diinterpretasikan bahwa data maupun model dalam penelitian ini reliabel.

5.2 Saran

Saran oleh penulis mengenai penelitian selanjutnya yaitu: menggunakan data penelitian dengan jumlah yang lebih besar apabila akan mengklasifikasikan data ke dalam lebih dari tiga kategori kelas; melakukan peningkatan pada *script* untuk *data preprocessing* agar mendapatkan fitur yang lebih baik sehingga dapat meningkatkan ketepatan klasifikasi; dan dapat dilakukan komparasi mengenai metode Multinomial *Naïve Bayes* dengan metode lain, seperti metode *Random Forest* atau *Logistic Regression* untuk mendapatkan hasil akurasi model klasifikasi yang lebih optimal.

DAFTAR PUSTAKA

- Abarna, S., Sheeba, J. I., Jayasrilakshmi, S., & Devaneyan, S. P. (2022). Identification of cyber harassment and intention of target users on social media platforms. *Engineering applications of artificial intelligence*, 115, 105283.
- Ahlida Nikmatul H., Didih R. C., & Christian S. K. Aditya. (2023). Classification of Bullying Comments on YouTube Streamer Comment Sections Using Naïve Bayes Classification. *Journal of Systems Engineering and Information Technology (JOSEIT)*, 2(1), 25-28.
- Akhter, A., Uzzal, K. A., & Polash, M. A. (2019). Cyber bullying detection and classification using multinomial *Naïve Bayes* and fuzzy logic. *Int. J. Math. Sci. Comput*, 5(4), 1-12.
- Amin, A. N., & Erihadiana, M. (2022). Pendidikan Multikultural dalam Perspektif QS. Ar Rum Ayat 22. *JHIP-Jurnal Ilmiah Ilmu Pendidikan*, 5(3), 895-900.
- Amin, K., Alfarauqi, M. D. A., & Khatimah, K. (2018). Social media, cyber hate, and racism. *Komuniti: Jurnal Komunikasi dan Teknologi Informasi*, 10(1), 3-10.
- Ash-Shidiq, M. A., & Pratama, A. R. (2021). Ujaran Kebencian Di Kalangan Pengguna Media Sosial Di Indonesia: Agama Dan Pandangan Politik. *AUTOMATA*, 2(1).
- Awan A., A., & Naviani A. (2023). *Learn how to build and evaluate a Naive bayes classifier using Python's Scikit-learn package*. Datacamp.
- Dihni, A. V. (2021). *Laporan Konten Kasus Kejahatan Siber (Januari-September 2021)*. Patrolisiber.
- Diyah, N. C. M. (2016). Kekerasan Dalam Pendidikan (Studi Fenomenologi Perilaku Kekerasan Di Panti Rehabilitasi Sosial Anak). *Paradigma*, 4(3).
- El Akbar, R. R., Shofa, R. N., & Paripurna, M. I. (2019, August). The implementation of Naïve Bayes algorithm for classifying tweets containing hate speech with political motive. In *2019 International Conference on Sustainable Engineering and Creative Computing (ICSECC)* (pp. 144-148). IEEE.
- Faisal, M. R., & Nugrahadi, D. T. (2020). Studi Ekstraksi Fitur Berbasis Vektor Word2Vec pada Pembentukan Fitur Berdimensi Rendah. *Jurnal Komputasi*, 8(1), 62-69.
- Feldman, R., & Sanger, J. (2007). *The text mining handbook: advanced approaches in analyzing unstructured data*. Cambridge university press.

- García, S., Luengo, J., & Herrera, F. (2015). *Data preprocessing in data mining*. Springer International Publishing Switzerland.
- Gorrell, G., Bakir, M. E., Roberts, I., Greenwood, M. A., & Bontcheva, K. (2020). Which politicians receive abuse? Four factors illuminated in the UK general election 2019. *EPJ data science*, 9(1), 18.
- Gorunescu, F. (2011). *Data Mining: Concepts, models and techniques* (Vol. 12). Springer Science & Business Media.
- Harahap, C. N., Marthasari, G. I., & Hayatin, N. (2021). Perbandingan Klasifikasi Berita Hoax Kategori Kesehatan Menggunakan *Naïve Bayes* dan Multinomial *Naïve Bayes*. *Jurnal Repositor*, 3(4), 419-424.
- Hosein, P., & Baboolal, K. (2022). *Bayes Classification using an approximation to the Joint Probability Distribution of the Attributes*. *arXiv*. Cornell University.
- Ibrahim, V., Bakar, J. A., Harun, N. H., & Abdulateef, A. F. (2021). A word cloud model based on hate speech in an online social media environment. *Baghdad Science Journal*, 18(2), 937-946.
- Jebreen, K. (2017). *Modèles graphiques pour la classification et les séries temporelles*. Aix-Marseille.
- Kango, U. (2009). Bentuk-Bentuk Kekerasan yang Dialami Perempuan. *Jurnal Legalitas*, 2(01).
- Kemenag. (2019). *Al-Qur'an dan Terjemahan*.
- Komnas Perempuan. (2022). Catatan Tahunan Kekerasan Terhadap Perempuan Tahun 2021.
- Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26). New York: Springer.
- Kurniati, K. (2022). Penegakan Hukum Terhadap Perilaku Tindak Pidana Pelecehan Seksual Melalui Media Sosial Dalam Pasal 27 Ayat 1 UU ITE Nomor 19 Tahun 2016. *Disiplin: Majalah Civitas Akademika Sekolah Tinggi Ilmu Hukum sumpah Pemuda*, 91-98.
- Kurniawan, T. (2017). Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Media Mainstream Menggunakan *Naïve Bayes Classifier* Dan Support Vector Machine. *Institut Teknologi Sepuluh Nopember*.

- Kusno, A. (2021). Redefinisi Kosakata Terkait Perang Bahasa (Kajian Linguistik Forensik). *Diglosia: Jurnal Kajian Bahasa, Sastra, dan Pengajarannya*, 4(3), 287-300.
- Latiefah, N. S., Sari, D. W., & Tarsidi, Z. D. (2022). Social Media Police to Shape Citizen Ethics in The Digital Age. *Jurnal Civicus*, 44-50.
- Muhammadun, M. (2011). Konsep Kejahatan Dalam Al-Quran (Perspektif Tafsir Maudhu'i). *DIKTUM: Jurnal Syariah dan Hukum*, 9(1), 14-29.
- Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.
- Mustika, A. L., Setiyono, S., Santoso, M., & Sabrina, N. (2021). Pertanggungjawaban Pidana Atas Tindak Pidana Pelecehan Verbal Melalui Media Sosial. *Bhirawa Law Journal*, 2(1), 67-72.
- Nurhafidzah, N. (2019). Analisis Topic Modeling dan Klasifikasi Ujaran Kebencian Menggunakan Algoritma Support Vector Machine. *Universitas Islam Indonesia*.
- Pardede, J., Miftahuddin, Y., & Kahar, W. (2020). Deteksi Komentar Cyberbullying Pada Media Sosial Berbahasa Inggris Menggunakan Naïve Bayes Classification. *Jurnal Informatika*, 7(1), 46-54.
- Prieto Curiel, R., Cresci, S., Muntean, C. I., & Bishop, S. R. (2020). Crime and its fear in social media. *Palgrave Communications*, 6(1).
- Putri, N. F. (2022). *Peran Lembaga Perlindungan Anak (LPA NTB) dalam menangani kasus kekerasan fisik terhadap anak di Mataram pada masa pandemi covid-19*. UIN Mataram.
- Putri, S. R., Ismaya, E. A., & Fardani, M. A. (2021). Fenomena Verbal Bullying Di Masyarakat Pedawang. *Jurnal Kajian Penelitian Dan Pendidikan Dan Pembelajaran*, 5(2), 792-796.
- Putri, T. A. M., Enri, U., & Sari, B. N. (2020). Analisis Algoritma *Naive bayes classifier* untuk Klasifikasi Tweet Pelecehan Seksual dengan #MeToo. *IJCIT (Indonesian J. Comput. Inf. Technol., vol. 5, no. 2*.
- Putri, W. T. H., & Hendrowati, R. (2018). Penggalan Teks Dengan Model Bag of Words Terhadap Data Twitter. *Jurnal Muara Sains, Teknologi, Kedokteran dan Ilmu Kesehatan*, 2(1), 129-138.
- Rahmawati, S. T. (2022). Kajian Kritis Terhadap Kejahatan Media Sosial Perspektif Al-Qur'an. *Jurnal Statement*, 55-67.
- Rahmi, K. M. (2021). Text Mining Analysis Dan Sentiment Analysis Dengan Menggunakan Metode *Naïve Bayes Classifier* (Studi Kasus: Data

Tanggapan Mengenai Tokopedia Melalui Media Sosial Twitter). *Universitas Islam Indonesia*.

- Rakhmawati, N. A., Aji, F., & Safira, R. (2021). Visualisasi Perundungan Pada *Live Streaming Gamer Wanita Menurut UU ITE 2008*. *Indonesian Journal of Business Intelligence (IJUBI)*, 3(2), 61-66.
- Reiten, T. E. G. (2017). *Classification with Multiple Classes using Naïve Bayes and Text Generation with a Small Data Set using a Recurrent Neural Network*. University of Agder.
- Rennie, J. D., Shih, L., Teevan, J., & Karger, D. R. (2003). Tackling the poor assumptions of naive bayes text classifiers. In *Proceedings of the 20th international conference on machine learning (ICML-03)* (pp. 616-623).
- Rinaldi, R., Goejantoro, R., & Syaripuddin. (2021). *Application of Naïve Bayes Multinomial Classification Method (Case Study: PT Prudential Life Samarinda in 2019)*. *Jurnal EKSPONENSIAL*, 12(2).
- Rizaty, A. M. (2022). *Pengguna Twitter di Indonesia Capai 18,45 Juta pada 2022*. DataIndonesia.
- Roifa, A. N. (2018). *Text Mining Dengan Metode Naive bayes classifier Untuk Mengklasifikasikan Berita Berdasarkan Konten*. Institut Teknologi Sepuluh Nopember.
- Russell, R. (2020). *Machine Learning: Step-by-Step Guide To Implement Machine Learning Algorithms with Python*. Knxb.
- Saputra, P. Y. (2017). Implementasi Teknik Crawling untuk Pengumpulan Data dari Media Sosial Twitter. *Dinamika Dotcom: Jurnal Pengembangan Manajemen Informatika dan Komputer*.
- Schutze, H., Manning, C. D., & Raghavan, P. (2008). *Introduction to information retrieval*. Cambridge University Press.
- Sembodo, J. E., Setiawan, E. B., & Baizal, Z. A. (2016, October). Data crawling otomatis pada twitter. In *Indonesian Symposium on Computing (Indo-SC)* (pp. 11-16).
- Southeast Asia Freedom of Expression Network. (2021). *Indonesia Digital Rights Situation Report 2020: Digital Repression Amid the Pandemic*.
- Stephen, A. (2022). *A good accuracy score in machine learning*. Sklearn metrics.
- Stevens, F., Nurse, J. R., & Arief, B. (2021). Cyber stalking, cyber harassment, and adult mental health: A systematic review. *Cyberpsychology, Behavior, and Social Networking*, 24(6), 367-376.

- Stone, J. V. (2013). *Bayes' Rule: A Tutorial Introduction to Bayesian Analysis*. Sebtel Press.
- Sugiyono. (2011). *Metode Penelitian Kuantitatif dan R&D Edisi Revisi*. Bandung: Alfabeta.
- Sunardi, S., Fadlil, A., & Kusuma, N. M. P. (2023). Comparing Data Mining Classification for Online Fraud Victim Profile in Indonesia. *INTENSIF: Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi*, 7(1), 1-17.
- Tirahmawan, J., Melody, B. A. L., & Ahly, M. N. N. (2021). Rasisme Terhadap Kulit Hitam dalam Iklan H&M. *Jurnal Audiens*, 2(1), 18-25.
- Twitter. (2016). Twitter Support. Diakses dari Twitter: [http:// support.twitter.com/](http://support.twitter.com/)
- Ulhaq, D., Suarna, N., & Dwilestari, G. (2022). Klasifikasi Berita Kriminal Menggunakan Algoritma Naïve Bayes Berbasis PSO. *INFORMATICS FOR EDUCATORS AND PROFESSIONAL: Journal of Informatics*, 6(2), 116-125.
- Vogels, E. A. (2021). The state of online harassment. *Pew Research Center*, 13, 625.
- Walpole, R. E. (1995). *Pengantar Statistika Edisi ke-3*. Jakarta: PT Gramedia Pustaka Utama.
- Widyaningsih, Y., Arum, G. P., & Prawira, K. (2021). Aplikasi K-Fold Cross Validation Dalam Penentuan Model Regresi Binomial Negatif Terbaik. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 15(2), 315-322.
- Winkelman, B., S., Oomen-Early, J., Walker, A. D., Chu, L., & Yick-Flanagan, A. (2015). Exploring cyber harassment among women who use social media. *Universal journal of public health*, 3(5), 194.
- Wolford-Clevenger, C., Zapor, H., Brasfield, H., Febres, J., Elmquist, J., Brem, M., ... & Stuart, G. L. (2016). An examination of the Partner Cyber Abuse Questionnaire in a college student sample. *Psychology of violence*, 6(1), 156.
- Zhai, C., & Massung, S. (2016). *Text data management and analysis: a practical introduction to information retrieval and text mining*. Morgan & Claypool Publishers.

LAMPIRAN

Lampiran 1. Data *Tweet* Kasus *Cyber Harassment*

No.	<i>Tweet</i>	Kelas
1.	@tanyakanrl nikah 3 kali dia main tangan??? verbally abusive?? doyan nampar anaknya??? doyan pentokin kepala anak ke kaca?? suka mukulin anaknya ampe ungu??? suka nyambit pake belt kulit???	<i>Physical Threats</i>
2.	@nurinazahr bye muka lu yg burik dan bopeng ituuu ih. belajar ye wahai anjing dan tolol biar bisa dpt kerja yg mapan 🤔🤔🤔	<i>Purposeful Embarrassment</i>
3.	@sukiman49 @sutanmangara Man sya mau tanya ada kah bukti yg lo sebut kadrun itu memperbudak WNI? Kalau Cina sipit banyak buktinya.	<i>Racist</i>
4.	Si ibu sudah lama gak dipake kali ya, makanya kayak anjing. Keroyok aja bagusnya, perkosa rame rame biar kapok!!! https://t.co/PIduBUV2FB	<i>Sexual harassment</i>
5.	@ArnoldPoernomo Its okay @Freya_JKT48 jangan dengarkan chef minoritas ini. Dasar tenaga kerja cina @ArnoldPoernomo	<i>Racist</i>
6.	@Leora_ceryllea 🌸 : rasanya darah saya bergelojak, langsung pingin saya bunuh. Harusnya tadi langsung ku bekap yah.. hmm aku tak kepikiran	<i>Physical threats</i>
7.	@Beritasatu Makanya dari kmrn Saya melakukan penentangan thd agama muhammad dajjal yg mengedepankan ritual ibadah.. Agama dajjal adalah Islam dan Agama Rabb adalah Muslim, sama seperti agama Ibrahim yaitu Muslim. Tdk ada ayat yg berkata bahwa Ibrahim adalah seorang islam, tp Muslim	<i>Racist</i>
8.	@Murniastutii Yah si hijaber tolol, yg kek gw aja gak mau sama hijaber tolol. Apalagi yg lebih Soleh dari gw, pasti lebih JIJK lagi 🤔🤔🤔	<i>Purposeful Embarrassment</i>
9.	@txtdrimedia Hukuman mati aja. Anak sendiri di perkosa berkali kali. Sampai hamil, anaknya dibunuh pula. Bejat banget sialan. Ibu anaknya (istri pelaku) juga ko bisa...	<i>Sexual Harassment</i>
10.	Pengen ku cabein mulut mertuanya. Pengen ku jambak rambut mertuanya Pengen ku tampol suaminya https://t.co/R3bldr3AuV	<i>Physical Threats</i>
11.	@Miduk17 ngapain KRISTEN IKUT CAMPUR URUSAN MUSLIM..???URUS SAJA AGAMAMU..DASAR KAFIR GA TAU DIRI..UDAH MINORITAS..NGLUNJAK...	<i>Racist</i>

12.	Kalau udah temen sefandomnya buat masalah, fandom orang dikambing hitamin, dasar kodok reog tolol. https://t.co/DpQYYzzhqK	<i>Purposeful Embarrassment</i>
13.	@ChoipanKucaai bacot cina	<i>Racist</i>
14.	@DeemasPermana @_Rizmaya__ @budimandjatmiko @aniesbaswedan Pantas cebong rata2 tolol bego2 lg.	<i>Purposeful embarrassment</i>
15.	TKA CINA ADALAH KEBIJAKAN PENGKHIANAT BANGSA! Jusuf Kalla: Indonesia Kaya Nikel, Tapi yang Kerja China Semua, Sampai ke Tukang Las https://t.co/8SG3nJL05m #LBPBrutusAntekChina #LBPBrutusAntekChina	<i>Racist</i>
16.	@FtmMybin @txtdrkuliner @gojekindonesia @GrabID Lah keluarga keluarga dia kenapa kamu yang kena ka, kenal aja kaga berani banget nampar dan menggunakan alasan yang tidak masuk di akal juga, bawa masalah keluarga ke pekerjaan apalagi sampe berani pake kekerasan ke orang lain hadehhh serem driver skrng 😬	<i>Physical Threats</i>
:	:	:
1.993	@BayuPur69164847 @jabarekspres Kebanyakan main game ini orang makanya tolol	<i>Purposeful Embarrassment</i>
1.994	@A_QuinnaNew2209 Panji gumilang ini memang anak anjing . PKI BIADAB panji gumilang ini .	<i>Purposeful Embarrassment</i>
1.995	@twinktwunktwonk @aeglioo Bunuh dia mas dom lempar pake besi panas	<i>Physical Threats</i>
1.996	@imjiiyeon Dia dibunuh dan air itu digunakan pembunuhnya untuk menghilangkan jejak. Kemudian mayatnya digantung seakan akan abis bunuh diri	<i>Physical Threats</i>
1.997	Lonte jayapura guys https://t.co/o9j4y0fdNE	<i>Sexual Harassment</i>
1.998	Si lonte stay hotel cendrawasih abadi entrop hati ² ya teman ² ini tukang penipu lonte kelas kakap di jayapura https://t.co/GUfXDV3KBC	<i>Sexual Harassment</i>
1.999	@Matahari_tmr Maka dari itu jangan sampai turunan Yaman berkuasa dan dipilih, kita itu sebenarnya lagi dijajah sama Orang timur tengah dengan memutar balikan fakta bahwa lagi dijajah sama Cina .	<i>Racist</i>
2.000	@NanunSanjaya @BBravo911 @P3nulis_alam1_ Dan lu adalah keturunan lonte yg diperkosa anjing buduk	<i>Sexual Harassment</i>

Lampiran 2. Hasil Data Preprocessing

No.	Case Folding	Cleansing	Tokenization	Normalization	Stopwords Removal	Stemming
1.	tanyakan rl nikah 3 kali dia main tangan?? ? verbally abusive? ? doyan nampar anaknya? ?? doyan pentokin kepala anak ke kaca suka anak ke kaca?? suka mukulin anaknya ampe ungu suka anaknya ampe ungu??? suka nyambit pake belt kulit???	tanyakanrl nikah kali dia main tangan verbally abusive doyan nampar anaknya doyan pentokin kepala anak ke kaca suka mukulin anaknya ampe ungu suka nyambit pake belt kulit	['tanyakanrl', , 'nikah', 'kali', 'dia', 'main', 'tangan', 'verbally', 'abusive', 'doyan', 'nampar', 'anaknya', 'doyan', 'pentokin', 'kepala', 'anak', 'ke', 'kaca', 'suka', 'mukulin', 'anaknya', 'ampe', 'anaknya', 'ampe', 'ungu', 'suka', 'nyambit', 'pake', 'belt', 'kulit']	['tanyakanrl', 'nikah', 'kali', 'dia', 'main', 'tangan', 'verbally', 'abusive', 'doyan', 'nampar', 'anaknya', 'doyan', 'pentokin', 'kepala', 'anak', 'ke', 'kaca', 'suka', 'mukulin', 'anaknya', 'ampe', 'ungu', 'suka', 'nyambit', 'pake', 'belt', 'kulit']	['tanyakan rl', 'nikah', 'kali', 'main', 'tangan', 'verbally', 'abusive', 'doyan', 'nampar', 'anaknya', 'doyan', 'pentokin', 'kepala', 'anak', 'kaca', 'suka', 'mukulin', 'anaknya', 'ampe', 'anaknya', 'ampe', 'ungu', 'suka', 'nyambit', 'pake', 'belt', 'kulit']	['tanyakan rl', 'nikah', 'kali', 'main', 'tangan', 'verbally', 'abusive', 'doyan', 'nampar', 'anak', 'doyan', 'pentokin', 'kepala', 'anak', 'kaca', 'suka', 'mukulin', 'anak', 'ampe', 'anak', 'ungu', 'suka', 'nyambit', 'pake', 'belt', 'kulit']
2.	@nurina zahr bye muka lu yg burik dan yg burik dan bopeng ituuu ih bopeng ituuu ih. belajar ye wahai belajar ye wahai anjing dan tolol biar bisa dpt kerja yg mapan yg mapan 👉👉 👉	bye muka lu yg burik dan bopeng ituuu ih belajar ye wahai anjing dan tolol biar bisa dpt kerja yg mapan	['bye', 'muka', 'lu', 'yg', 'burik', 'dan', 'bopeng', 'ituuu', 'ih', 'belajar', 'ye', 'wahai', 'anjing', 'dan', 'tolol', 'biar', 'bisa', 'dpt', 'kerja', 'yg', 'mapan']	['bye', 'muka', 'lu', 'yg', 'burik', 'dan', 'bopeng', 'ituuu', 'ih', 'belajar', 'ye', 'wahai', 'anjing', 'dan', 'tolol', 'biar', 'bisa', 'dpt', 'kerja', 'yg', 'mapan']	['bye', 'muka', 'burik', 'bopeng', 'ituuu', 'ih', 'belajar', 'ye', 'anjing', 'tolol', 'biar', 'dpt', 'kerja', 'mapan']	['bye', 'muka', 'burik', 'bopeng', 'ituuu', 'ih', 'ajar', 'ye', 'anjing', 'tolol', 'biar', 'dpt', 'kerja', 'mapan']
3.	@sukima n49	man sya mau tanya	['man', 'sya', 'mau',	['man', 'sya', 'mau', 'tanya',	['man', 'sya', 'kah',	['man', 'sya', 'kah',

	@sutanman angara man sya mau tanya ada kah bukti yg lo sebut kadrun itu memperbudak wni kalau cina sipit banyak itu memperbudak wni? kalau cina sipit banyak buktinya.	ada kah bukti yg lo sebut kadrun itu memperbudak wni kalau cina sipit banyak buktinya	'tanya', 'ada', 'kah', 'bukti', 'yg', 'lo', 'sebut', 'kadrun', 'itu', 'memperbudak', 'wni', 'kalau', 'cina', 'sipit', 'banyak', 'buktinya']	'ada', 'kah', 'bukti', 'yg', 'lo', 'sebut', 'kadrun', 'itu', 'memperbudak', 'wni', 'kalau', 'cina', 'sipit', 'banyak', 'buktinya']	'bukti', 'kadrun', 'memperbudak', 'wni', 'cina', 'sipit', 'buktinya']	'bukti', 'kadrun', 'budak', 'wni', 'cina', 'sipit', 'bukti']
4.	si ibu sudah lama gak dipake kali ya, makanya kayak anjing. keroyok aja bagusnya , perkosa rame rame biar kapok!!! https://t.co/pidubuv2fb	si ibu sudah lama gak dipake kali ya makanya kayak anjing keroyok aja bagusnya perkosa rame rame biar kapok	['si', 'ibu', 'sudah', 'lama', 'gak', 'dipake', 'kali', 'ya', 'makanya', 'kayak', 'anjing', 'keroyok', 'aja', 'bagusnya', 'perkosa', 'rame', 'rame', 'biar', 'kapok']	['si', 'ibu', 'sudah', 'lama', 'gak', 'dipake', 'kali', 'ya', 'makanya', 'kayak', 'anjing', 'keroyok', 'aja', 'bagusnya', 'perkosa', 'rame', 'rame', 'biar', 'kapok']	['dipake', 'kali', 'kayak', 'anjing', 'keroyok', 'bagusnya', 'perkosa', 'rame', 'rame', 'biar', 'kapok']	['dipake', 'kali', 'kayak', 'anjing', 'keroyok', 'bagus', 'perkosa', 'rame', 'rame', 'biar', 'kapok']
5.	@arnoldpoernom o its okay @freya_jkt48 jangan dengarkan chef minoritas ini. dasar tenaga kerja	its okay jkt jangan dengarkan chef minoritas ini dasar tenaga kerja cina	['its', 'okay', 'jkt', 'jangan', 'dengarkan', 'chef', 'minoritas', 'ini', 'dasar', 'tenaga', 'kerja', 'cina']	['its', 'okay', 'jkt', 'jangan', 'dengarkan', 'chef', 'minoritas', 'ini', 'dasar', 'tenaga', 'kerja', 'cina']	['its', 'okay', 'jkt', 'dengarkan', 'chef', 'minoritas', 'tenaga', 'kerja', 'cina']	['its', 'okay', 'jkt', 'dengar', 'chef', 'minoritas', 'tenaga', 'kerja', 'cina']

	cina @arnold poernom o					
6.	@leora_ ceryllea 🌀: rasanya darah saya bergeloka, langsung pingin saya bunuh harusnya tadi langsung ku bekap yah hmm tadi langsung ku bekap yah.. hmm aku tak kepikiran	ceryllea rasanya darah saya bergelajak langsung pingin saya bunuh harusnya tadi langsung ku bekap yah hmm aku tak kepikiran	['ceryllea', 'rasanya', 'darah', 'saya', 'bergelajak', 'langsung', 'pingin', 'saya', 'bunuh', 'harusnya', 'tadi', 'langsung', 'ku', 'bekap', 'yah', 'hmm', 'aku', 'tak', 'kepikiran']	['ceryllea', 'rasanya', 'darah', 'saya', 'bergelajak', 'langsung', 'pingin', 'saya', 'bunuh', 'harusnya', 'tadi', 'langsung', 'ku', 'bekap', 'yah', 'hmm', 'aku', 'tak', 'kepikiran']	['ceryllea', 'darah', 'bergelajak', 'langsung', 'pingin', 'langsung', 'ku', 'bekap', 'yah', 'hmm', 'kepikiran']	['ceryllea', 'darah', 'bergelajak', 'langsung', 'pingin', 'langsung', 'ku', 'bekap', 'yah', 'hmm', 'pikir']
∴	∴	∴	∴	∴	∴	∴
1.998	Si lonte stay hotel cendrawa sih abadi entrop hati ² ya teman ² ini tukang penipu lonte kelas kakap di jayapura https://t.co o/GUfX DV3KB C	si lonte stay hotel cendrawasih abadi entrop hati ² ya teman ² ini tukang penipu lonte kelas kakap di jayapura https://t.co /gufxdv3k bc	si lonte stay hotel cendrawasih abadi entrop hati ya teman ini tukang penipu lonte kelas kakap di jayapura	['si', 'lonte', 'stay', 'hotel', 'cendrawasih', 'abadi', 'entrop', 'hati', 'ya', 'teman', 'ini', 'tukang', 'penipu', 'lonte', 'kelas', 'kakap', 'di', 'jayapura']	['si', 'lonte', 'stay', 'hotel', 'cendrawasih', 'abadi', 'entrop', 'hati', 'ya', 'teman', 'ini', 'tukang', 'penipu', 'lonte', 'kelas', 'kakap', 'di', 'jayapura']	['lonte', 'stay', 'hotel', 'cendrawasih', 'abadi', 'entrop', 'hati', 'teman', 'tukang', 'penipu', 'lonte', 'kelas', 'kakap', 'jayapura']
1.999	@Matahari_tmr Maka dari itu jangan	@matahari_tmr maka dari itu jangan sampai	tmr maka dari itu jangan sampai turunan	['tmr', 'maka', 'dari', 'itu', 'jangan', 'sampai', 'turunan',	['tmr', 'maka', 'dari', 'itu', 'jangan', 'sampai',	['tmr', 'turunan', 'yaman', 'berkuasa', 'dipilih',

	sampai turunan Yaman berkuasa dan dipilih, kita itu sebenarnya lagi dijajah sama Orang timur tengah dengan memutar balikan fakta bahwa lagi dijajah sama Cina .	turunan yaman berkuasa dan dipilih, kita itu sebenarnya lagi dijajah sama orang timur tengah dengan memutar balikan fakta bahwa lagi dijajah sama cina .	yaman berkuasa dan dipilih kita itu sebenarnya lagi dijajah sama orang timur tengah dengan memutar balikan fakta bahwa lagi dijajah sama cina	'yaman', 'berkuasa', 'dan', 'dipilih', 'kita', 'itu', 'sebenarnya', 'lagi', 'dijajah', 'sama', 'orang', 'timur', 'tengah', 'dengan', 'memutar', 'balikan', 'fakta', 'bahwa', 'lagi', 'dijajah', 'sama', 'cina']	'turunan', 'yaman', 'berkuasa', 'dan', 'dipilih', 'kita', 'itu', 'sebenarnya', 'lagi', 'dijajah', 'sama', 'orang', 'timur', 'tengah', 'dengan', 'memutar', 'balikan', 'fakta', 'bahwa', 'lagi', 'dijajah', 'sama', 'cina']	'dijajah', 'timur', 'memutar', 'balikan', 'fakta', 'dijajah', 'cina']
2.000	@nanuns anjaya @bbravo 911 @p3nulis_alam1_ dan lu adalah keturunan lonte yg diperkosa anjing buduk	alam dan lu adalah keturunan lonte yg diperkosa anjing buduk	['alam', 'dan', 'lu', 'adalah', 'keturunan', 'lonte', 'yg', 'diperkosa', 'anjing', 'buduk']	['alam', 'dan', 'lu', 'adalah', 'keturunan', 'lonte', 'yg', 'diperkosa', 'anjing', 'buduk']	['lu', 'adalah', 'keturunan', 'lonte', 'diperkosa', 'anjing', 'buduk']	['lu', 'adalah', 'turun', 'lonte', 'perkosa', 'anjing', 'buduk']

Lampiran 3. Hasil Ekstraksi Fitur

Dok.	Kata												
	1.	...	183	184	...	668	669	670	...	1.466	1.467	1.468	...
	abad	...	anak	anarkis	...	cina	cinta	cipta	...	kaffah	kafir	kaisar	...
1.	0	...	3	0	...	0	0	0	...	0	0	0	...
2.	0	...	0	0	...	0	0	0	...	0	0	0	...
3.	0	...	0	0	...	1	0	0	...	0	0	0	...
4.	0	...	0	0	...	0	0	0	...	0	0	0	...
5.	0	...	0	0	...	1	0	0	...	0	0	0	...
6.	0	...	0	0	...	0	0	0	...	0	0	0	...
7.	0	...	0	0	...	0	0	0	...	0	0	0	...
8.	0	...	0	0	...	0	0	0	...	0	0	0	...
9.	0	...	3	0	...	0	0	0	...	0	0	0	...
10.	0	...	0	0	...	0	0	0	...	0	0	0	...
11.	0	...	0	0	...	0	0	0	...	0	1	0	...
12.	0	...	0	0	...	0	0	0	...	0	0	0	...
13.	0	...	0	0	...	1	0	0	...	0	0	0	...
14.	0	...	0	0	...	0	0	0	...	0	0	0	...
15.	0	...	0	0	...	1	0	0	...	0	0	0	...
16.	0	...	0	0	...	0	0	0	...	0	0	0	...
17.	0	...	0	0	...	0	0	0	...	0	0	0	...
18.	0	...	0	0	...	0	0	0	...	0	0	0	...
19.	0	...	0	0	...	0	0	0	...	0	0	0	...
20.	0	...	0	0	...	2	0	0	...	0	0	0	...
...
1.987	0	...	0	0	...	0	0	0	...	0	0	0	...
1.988	0	...	0	0	...	0	0	0	...	0	1	0	...
1.989	0	...	0	0	...	0	0	0	...	0	0	0	...
1.990	0	...	0	0	...	0	0	0	...	0	0	0	...
1.991	0	...	0	0	...	0	0	0	...	0	0	0	...
1.992	0	...	0	0	...	0	0	0	...	0	0	0	...
1.993	0	...	0	0	...	0	0	0	...	0	0	0	...
1.994	0	...	1	0	...	0	0	0	...	0	0	0	...
1.995	0	...	0	0	...	0	0	0	...	0	0	0	...
1.996	0	...	0	0	...	0	0	0	...	0	0	0	...
1.997	0	...	0	0	...	0	0	0	...	0	0	0	...
1.998	0	...	0	0	...	0	0	0	...	0	0	0	...
1.999	0	...	0	0	...	1	0	0	...	0	0	0	...
2.000	0	...	0	0	...	0	0	0	...	0	0	0	...

Dok.	Kata										
	2.407	2.408	...	3.276	3.277	3.278	...	4.082	4.083	...	5.728
	muslim	musnah	...	rambut	rame	rampok	...	tolol	tolong	...	zuckerberg
1.	0	0	...	0	0	0	...	0	0	...	0
2.	0	0	...	0	0	0	...	1	0	...	0
3.	0	0	...	0	0	0	...	0	0	...	0
4.	0	0	...	0	2	0	...	0	0	...	0
5.	0	0	...	0	0	0	...	0	0	...	0
6.	0	0	...	0	0	0	...	0	0	...	0
7.	3	0	...	0	0	0	...	0	0	...	0
8.	0	0	...	0	0	0	...	2	0	...	0
9.	0	0	...	0	0	0	...	0	0	...	0
10.	0	0	...	1	0	0	...	0	0	...	0
11.	1	0	...	0	0	0	...	0	0	...	0
12.	0	0	...	0	0	0	...	1	0	...	0
13.	0	0	...	0	0	0	...	0	0	...	0
14.	0	0	...	0	0	0	...	1	0	...	0
15.	0	0	...	0	0	0	...	0	0	...	0
16.	0	0	...	0	0	0	...	0	0	...	0
17.	0	0	...	0	0	0	...	0	0	...	0
18.	0	0	...	0	0	0	...	1	0	...	0
19.	0	0	...	0	0	0	...	0	0	...	0
20.	0	0	...	0	0	0	...	0	0	...	0
...
1.987	0	0	...	0	0	0	...	0	0	...	0
1.988	0	0	...	0	0	0	...	0	0	...	0
1.989	0	0	...	0	0	0	...	0	0	...	0
1.990	0	0	...	0	0	0	...	0	0	...	0
1.991	0	0	...	0	0	0	...	0	0	...	0
1.992	0	0	...	0	0	0	...	0	0	...	0
1.993	0	0	...	0	0	0	...	1	0	...	0
1.994	0	0	...	0	0	0	...	0	0	...	0
1.995	0	0	...	0	0	0	...	0	0	...	0
1.996	0	0	...	0	0	0	...	0	0	...	0
1.997	0	0	...	0	0	0	...	0	0	...	0
1.998	0	0	...	0	0	0	...	0	0	...	0
1.999	0	0	...	0	0	0	...	0	0	...	0
2.000	0	0	...	0	0	0	...	0	0	...	0

Lampiran 4. Hasil Klasifikasi Jenis-Jenis *Cyber Harassment* Menggunakan *Multinomial Naïve Bayes*

No.	Kalimat	Hasil Klasifikasi
1.	bunuh orang kek gtu	<i>Racist</i>
2.	dih anjing babi bangsat ngentot	<i>Purposeful Embarrassment</i>
3.	cina sepet ngko hina cipta allah swt ye perangai rasis yahudi zionis islam tidak ajar hina orang surah al hujurat maktub allah swt cipta berbagai kaum	<i>Racist</i>
4.	token goblok tim tolol lonte lacur busuk bangsaaad anjing babi scam token	<i>Purposeful Embarrassment</i>
5.	sanusi tolol gak ada obat	<i>Purposeful Embarrassment</i>
6.	redsindo fans club tolol hiburan diri	<i>Purposeful Embarrassment</i>
7.	panji gumilang memang anak anjing pki biadab panji gumilang	<i>Racist</i>
8.	bangsat lonte	<i>Sexual Harassment</i>
9.	Faaakkkk gua mau bunuh orang rasa	<i>Physical Threats</i>
10.	iya mereka bener ngancam ancaman kosong doang pasca jadi warga balas dendam cerita ibu tokoh pki seret massa keliling desa krn bener simbol besar gila kuasa korup jatah beras maen ancaman dll	<i>Racist</i>
11.	jujurrrr rasa pengen pukul	<i>Physical Threats</i>
12.	lacur rupany	<i>Sexual Harassment</i>
13.	otak tolol kalian gak nyampe dgn maksud beliau baris kaum ikut dajjal kayak kalian tolol	<i>Purposeful Embarrassment</i>
14.	wibu cabul tolol idiot	<i>Purposeful Embarrassment</i>
15.	bacot lonte	<i>Sexual Harassment</i>
16.	pesilat nyali patungan goblok sumpah	<i>Purposeful Embarrassment</i>
17.	Emak bapak gua manggil lonte gara lembur kerja pengen ganti ortu	<i>Physical Threats</i>
18.	bokap lu langgan lonte pinggir rel	<i>Sexual Harassment</i>
19.	bangsat anjing goblok babi liar binatang mmk ktl ngtt biadab sundal	<i>Racist</i>
20.	tka cina bijak khianat bangsa jusuf kalla indonesia kaya nikel kerja china tukang las	<i>Racist</i>
21.	kaum munafiqun zein kribo anjing buzzer	<i>Purposeful Embarrassment</i>

22.	bacot puja bencong homo lonte lonte mek sange kali nonton video bencong nyanyi najiss	<i>Sexual Harassment</i>
23.	bodoh kali bilang rahasia dispill gc goblok kali sumpah	<i>Purposeful Embarrassment</i>
24.	emaknya lonte	<i>Sexual Harassment</i>
25.	kalo mba pukul mau	<i>Physical Threats</i>
26.	tv one ikut orgil stres jg spt israel ma yahudi ngundang narsum kualitas sok toleransi pdhl asli preeett	<i>Racist</i>
27.	jokowi benci islam	<i>Racist</i>
28.	anda pikir cina kaya krn siapa naga dapet modal darimana orang percaya soeharto cina kroni soeharto jaman orba uang negara ngalir keluarga cendana liat list blbi kl tdk percaya	<i>Racist</i>
29.	bilang ngatain belle lonte yakan admin berani ngasi wadah biar audiens ngatain lonte trus cuci tangan pura innocent	<i>Sexual Harassment</i>
30.	keturunan cina rata bajingan	<i>Racist</i>
31.	ngeselin banget orang pengen bunuh gua mutilasi abis bakar abu ludahin campur tai kucing	<i>Physical Threats</i>
32.	pukul pake gantung jemur el	<i>Physical Threats</i>
33.	blom liat cctvnya gk ppercaya ramen narasi terlalu gebu bela ojol usut mukul langsung tindak pidana ramen lapor pake pasal karet doxxingnya	<i>Physical Threats</i>
34.	Moga orang gw benci bisa gw pukul orang ajar kali	<i>Physical Threats</i>
35.	cewe sange cowo pacar ade kakak mama tante pacar kirim pap dm ntar lecehin privasi aman pejuh	<i>Sexual Harassment</i>
36.	cina omong besar pencitraan	<i>Racist</i>
37.	gua pukul lu wan	<i>Physical Threats</i>
38.	uang pusat bebas pemda dki blok goblok itu wilayah dki lahan diskusi kpda rakyat tinggal bayar tolol km blok	<i>Purposeful Embarrassment</i>
39.	tetep bingung gimana pukul ya	<i>Physical Threats</i>
40.	kalo gak dosa serius gua bunuh	<i>Physical Threats</i>
41.	mgk hasil hub larang tinggal cwo nghamin kasi lahir dr cwe biadab spt pirang anjing nazis mugholadoh cuuihh berani sm bocah taik lonte	<i>Sexual Harassment</i>
42.	bapak tolol	<i>Purposeful Embarrassment</i>
43.	thread lonte lewat pagi pagi cok muntah	<i>Sexual Harassment</i>
44.	sabar tahun sikap anjing gajelas goblok tolol egois menang nerima fakta	<i>Purposeful Embarrassment</i>
45.	kesel anjing tag fuck you hiling chanel gerard nongol bapa bapa kumur kumur	<i>Physical Threats</i>

46.	gaya elu mau gabung mereka jakarta medan tolol begok lihat	<i>Purposeful Embarrassment</i>
47.	lu klo bs gwe bunuh	<i>Physical Threats</i>
48.	coba memek kedut mulu mending jadi lonte	<i>Sexual Harassment</i>
49.	lonte baru nonton bola diem	<i>Sexual Harassment</i>
50.	pukul guru pukull	<i>Physical Threats</i>
51.	dasar lonte udha sakit gak punya fogure bapak lu colong bapak orang lu lonte lonte	<i>Sexual Harassment</i>
52.	pengen nendang muka	<i>Physical Threats</i>
53.	jelek bahasa lu bahasa bekas	<i>Racist</i>
54.	tolol kali dia	<i>Purposeful Embarrassment</i>
55.	w pukul ye physical attack kaya	<i>Physical Threats</i>
56.	waduhh sekolah negeri spesifikasi utk guru muslim ngajar bahasa kafir colek dki kemdikbud negara pancasila negara agama	<i>Racist</i>
57.	pengen tak jambak sampek lepas rambut ci	<i>Physical Threats</i>
58.	online om kl cs duluan mukul gmn cm sabar gt	<i>Physical Threats</i>
59.	jambak aja kak	<i>Physical Threats</i>
60.	agama hina klo ibu hina aq marah klo kmu marah karn lonte jd ngerti milik sayang marah kmu cinta hina dina sy paham kmu ngerti krn kmu lonte	<i>Sexual Harassment</i>
61.	saya tdk butuh pandang yahudi tp sebagai pengguna kitab yahudi bawa saya data quran ibrahim mekkah dalih kanan kiri tertawa pandang yahudi mati kutu	<i>Racist</i>
62.	pgn pukul pala lu asu	<i>Physical Threats</i>
63.	takut anjeng bangsat brengsek tai sampah kayak gini hancurin hidup orang sadar woiiii lonte ngaku pria ngajak tidur biar dapet peran pete situ udh buka aib klau lu sampah	<i>Sexual Harassment</i>
64.	tweet hoax kadrun teroris kntl tukang fitnah dicariin polisi anjing	<i>Purposeful Embarrassment</i>
65.	bacot yahudi	<i>Racist</i>
66.	pribumi goblok perjuang non pribumi uda non pribumi niat busuk tidak mampu mkin cari makan dgn gtu kasihan	<i>Racist</i>
67.	gitu botak cina rasis baik borgol aja buzzer	<i>Racist</i>
68.	ahok baru jadi wni tahun warga cina darat siapa pribumi siapa ngaseng	<i>Racist</i>
69.	emg tampang gembel gmn diliatin gt pgn tampar pake duit	<i>Physical Threats</i>
70.	lu kontol otak lu anjing banyak ads anak lonte	<i>Sexual Harassment</i>

71.	klo goblok dari lahir bang percuma dibuktiin kaya apah percuma yah nama ngejilat daki selangkang heru trus puji puji gubernur skarang	<i>Physical Threats</i>
72.	mama kena kdrt teriak nyuruh bapak jangan mukul oppung boru dr bapak nyuruh diam sakit hati sampai sekarang meski bapak lebih baik dn oppung borunya meninggal	<i>Physical Threats</i>
73.	bagus hasil baik goblok kayak wan jembut harap korupsi bansos asu	<i>Purposeful Embarrassment</i>
74.	lu anjng kafir mafia chino ahokeer gerombol garong duit rakyat indonesia jiaah chino malu ngakui rrc chino komunis	<i>Racist</i>
75.	katain lonte pasang sendiri rasa ajir bngt	<i>Sexual Harassment</i>
76.	sodok mantap dri belakng jambak	<i>Physical Threats</i>
77.	klub tolol	<i>Purposeful Embarrassment</i>
78.	cerita dongeng makany belajar islam lu pelajari dongeng bukan aseli wahyu tuhan dongeng gabriel jibril islam ma michael mikhail islam perang lucu ada malaikat perang hadeehhh ancur dunia	<i>Racist</i>
79.	hancur panji gumilang hancur yahudi israel hancur masyarakat sipil bodoh hancur imam	<i>Racist</i>
80.	buku syariat islam muslim boleh fitnah agama lain tidak boleh sebaliknya kata jujur tentang islam keliatan belang fitnah tdk guna	<i>Racist</i>
81.	loe salah lonte muara angke bnyk lonte ngomong pikir pake otak liat pake mata bener lonte ngatain jimmin dance pake baju sopan blog ngakak sedih hidup dengki	<i>Sexual Harassment</i>
82.	ndak jauh hp juta umat buat cina bego piara yaaaa	<i>Racist</i>
83.	halah lonte pulsa kontol	<i>Sexual Harassment</i>
84.	lonte kayak gini entot mau	<i>Sexual Harassment</i>
85.	mana bisa kristen tanpa tanakh awal mula manusia kenal kenal allah islam tanpa taurat dapat ide darimana nabi beda islam spt kacang lupa kulit anggap ajar paling revolusioner padahal hal baru dalam jack halu	<i>Racist</i>
86.	coba dia perkosa	<i>Sexual Harassment</i>
87.	pengen pukul pala	<i>Physical Threats</i>
88.	kakak muslim kafir darah ulama racun	<i>Racist</i>
89.	ajaran islam sendiri boleh nabi homy trus	<i>Racist</i>
90.	syiet gua pukul lu	<i>Physical Threats</i>
91.	aing kata muka cina gimana bang minoritas	<i>Racist</i>

92.	cina parasit parasit ekonomi	<i>Racist</i>
93.	lonte puja bencong kayak lonte temen lu	<i>Sexual Harassment</i>
94.	ngomong anies kedengar anjing sich	<i>Purposeful Embarrassment</i>
95.	islam is haram	<i>Racist</i>
96.	yahudi ribu yahudi mati tak usik juta yahudi dunia banding muslim saudara muslim real fact satu tubuh	<i>Racist</i>
97.	daki yahudi	<i>Racist</i>
98.	masalahe lonte bcs iku ngrusuhi tim tok supporter barang yo lok lokno wes vakum dukung arema yo jik aremania	<i>Sexual Harassment</i>
99.	kadrun tolol kl ngehoak makan tahun drun	<i>Purposeful Embarrassment</i>
100.	Pengen cabe mulut mertua pengen jambak rambut mertua pengen tampol suami	<i>Physical Threats</i>

Lampiran 5. Script Data Crawling

```

#import library
library("rtweet")
library("dplyr")
library("tidyr")
library("writexl")

keyword = Bunuh OR pukul OR mukul OR tampar OR nampar OR jambak OR
tending OR nendang OR bodoh OR bego OR tolol OR goblok OR asu OR
anjing OR sampah OR pecundang OR pengecut OR jelek OR idiot OR islam
OR cina OR muslim OR masjid OR yahudi OR hitam OR putih OR kafir OR
perkosa OR sentuh OR raba OR cat calling OR pelecehan OR leceh OR jalang
OR lonte OR pelacur
tottweet = 1200, 1000, 1500, 800, 1000, 500, 1000, 1000, 500
type="recent"
bahasa="id"
retweet = FALSE

twitter_token <- create_token(app="SKRIPSI",
                             consumer_key="",
                             consumer_secret=" ",
                             access_token=" ",
                             access_secret="")

crawling_dataraw = search_tweets(keyword,
                                n = tottweet,
                                include_rts = retweet,
                                type = type,
                                lang = bahasa,
                                retryonratelimit = FALSE)

View(crawling_dataraw)

# select/filtering data
df_dataraw <- subset(crawling_dataraw, select=c(created_at, full_text))
datarawfinal <- na.exclude(df_dataraw)

write_xlsx(datarawfinal, "datarawfinal.xlsx")

```

Lampiran 6. Script Data Preprocessing

```

import pandas as pd
import numpy as np
import nltk
import argparse, os, re, string, pickle
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
from wordcloud import WordCloud
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings("ignore")

tweet_df = pd.read_excel('/content/datalabelling_final.xlsx')

df = tweet_df
df.head()

df.drop_duplicates(inplace=True)

df = pd.DataFrame(df[['tweet']])
df

df['Case Folding'] = df['tweet'].str.lower()
df

#Data Cleansing
#Remove Special Char
def remove_tweet(text):
    #remove tab, new line, and back slice
    text = str(text).replace('\t', " ").replace('\n', " ").replace('\u', " ").replace('\\', "")
    #remove non ASCII (emoticon, etc)
    text = text.encode('ascii', 'replace').decode('ascii')
    text = re.sub(r'^rt[\s]+', "", text)
    #remove mention, link, hashtag
    text = ''.join(re.sub("(@[A-Za-z0-9]+)|(^0-9A-Za-z \t)|(\w+:\w+\S+)", "", text).split())
    #remove incomplete URL
    return text.replace("http://", " ").replace("https://", " ")
df['Data Cleansing']=df['Case Folding'].apply(remove_tweet)
df.head()

#remove number
def remove_number(text):
    return re.sub(r"\d+", "", text)
df['Data Cleansing']=df['Data Cleansing'].apply(remove_number)

#remove user
def remove_pattern(input_txt, pattern):
    r = re.findall(pattern, input_txt)
    for i in r:
        input_txt = re.sub(i, "", input_txt)
    return input_txt

```

```

df['Data Cleansing'] = np.vectorize(remove_pattern)(df['Data Cleansing'], "@[\w]*")

#remove single char
def remove_single_char(text):
    return re.sub('\b[A-Za-z] \b|\b [A-Za-z]\b', "", text)
df['Data Cleansing']=df['Data Cleansing'].apply(remove_single_char)
df.head()

#Tokenization
def tokenization(text):
    text = re.split('\W+', text)
    return text
df['Tokenisasi']= df['Data Cleansing'].apply(tokenization)
df

df.to_csv('/content/hasilnormalisasi.csv')

#Normalisasi
normalized_word = pd.read_csv('hasilnormalisasi.csv')

normalized_word_dict = {}

for index, row in normalized_word.iterrows():
    if row[0] not in normalized_word_dict:
        normalized_word_dict[row[0]] = row[1]

def normalized_term(document):
    return [normalized_word_dict[term] if term in normalized_word_dict else term for
    term in document]

df['Normalisasi'] = df['Tokenisasi'].apply(normalized_term)
df

nltk.download('stopwords')

#Stopwords Removal
stopword = nltk.corpus.stopwords.words('indonesian')
stopword.extend([])
stopword= set(stopword)

def stopwords(text):
    text= [word for word in text if word not in stopword]
    return text

df["Stopwords"]= df["Normalisasi"].apply(stopwords)
df.head()

#Stemming
factory = StemmerFactory()
stem_say = factory.create_stemmer()
def stemmed_wrapper(term):
    return stem_say.stem(term)
term_base = {}

```

```
for document in df['Stopwords']:
    for term in document:
        if term not in term_base:
            term_base[term] = " "

print(len(term_base))
print("-----")

for term in term_base:
    term_base[term] = stemmed_wrapper(term)
    print(term,":",term_base[term])

print(term_base)
print("-----")

def get_stemmed_term(document):
    return [term_base[term] for term in document]

df['Stemmer'] = df['Stopwords'].apply(get_stemmed_term)

print(df['Stemmer'])

df.to_excel('/content/datahasil_preprocess.xlsx')
```

Lampiran 7. Script Multinomial Naïve Bayes

```
from sklearn.naive_bayes import MultinomialNB
model = MultinomialNB()
model.fit(x_trainvect,y_train)

predictions_mnb = model.predict(x_testvect)

print('\nClassification report:')
cr_MNB = classification_report(y_test, predictions_mnb)
print(cr_MNB)

df_mnbhasil = pd.DataFrame(predictions_mnb)
df_mnbhasil.head()

df_mnbhasil.to_excel('hasilprediksi.xlsx',index=False)

from sklearn import metrics
conf_mat = metrics.confusion_matrix(y_test, predictions_mnb)
class_label = ['physical threats','purposeful embarrassment','racist','sexual harassment']
train = pd.DataFrame(conf_mat, index = class_label, columns = class_label)
sns.heatmap(train, annot = True,fmt="d",cmap='Blues')
plt.title("Confusion Matrix")
plt.xlabel("Predicted Label")
plt.ylabel("True Label")
plt.show()
```

RIWAYAT HIDUP



Kiki Oktavia, lahir di Jakarta pada 26 Oktober 2000. Penulis merupakan anak terakhir dari empat bersaudara dari pasangan Bapak Saut dan Ibu Prahyati. Penulis telah menempuh pendidikan mulai dari TK RA Tarbiyatul Hidayah dan lulus pada tahun 2007, dilanjutkan menempuh pendidikan sekolah dasar di SDN Mangga Dua 01 Pagi dan lulus pada tahun 2013. Kemudian penulis melanjutkan pendidikan sekolah menengah pertama di SMPN 34 Jakarta dan lulus pada tahun 2016. Selanjutnya menempuh pendidikan sekolah menengah atas di SMAN 40 Jakarta dan lulus pada tahun 2019. Pada tahun yang sama, penulis melanjutkan pendidikan di Universitas Islam Negeri Maulana Malik Ibrahim Malang pada Program Studi Matematika, Fakultas Sains dan Teknologi.

Selama menempuh pendidikan di perguruan tinggi, penulis aktif mengikuti beberapa kegiatan baik di dalam lingkup kampus maupun di luar kampus. Penulis bergabung dalam organisasi mahasiswa yaitu Himpunan Mahasiswa Jurusan (HMJ) “Integral” Matematika selama satu periode sebagai anggota dan koordinator pada divisi Penerbitan dan Jurnalistik. Penulis juga bergabung dalam komunitas Literasi Matematika Integral (LIMIT) sebagai anggota dan ketua bagian Redaksional selama satu periode. Selain itu, penulis juga mengikuti berbagai kegiatan di luar kampus seperti aktivitas relawan, pelatihan, dan seminar.



KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI
MAULANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI

Jl. Gajayana No.50 Dinoyo Malang Telp. / Fax. (0341)558933

BUKTI KONSULTASI SKRIPSI

Nama : Kiki Oktavia
NIM : 19610079
Fakultas/Program Studi : Sains dan Teknologi/Matematika
Judul Skripsi : Pendekatan Multinomial *Naive Bayes* untuk Klasifikasi Jenis-Jenis *Cyber Harassment* di Media Sosial Twitter
Pembimbing I : Ria Dhea Layla N. K., M.Si.
Pembimbing II : Ari Kusumastuti, M.Pd., M.Si.

No	Tanggal	Hal	Tanda Tangan
1.	21 Desember 2022	Konsultasi Judul dan Topik	1.
2.	2 Februari 2023	Konsultasi BAB I dan BAB II	2.
3.	13 Februari 2023	Konsultasi Revisi BAB I	3.
4.	17 Mei 2023	Konsultasi Revisi BAB II	4.
5.	24 Mei 2023	Konsultasi Revisi BAB II dan BAB III	5.
6.	12 Juni 2023	Konsultasi Keagamaan BAB I dan BAB II	6.
7.	14 Juni 2023	Konsultasi Revisi BAB II dan BAB III	7.
8.	15 Juni 2023	ACC Proposal Skripsi	8.
9.	16 Juni 2023	Konsultasi Revisi dan ACC Keagamaan BAB I dan BAB II	9.
10.	13 Juli 2023	Konsultasi Keagamaan BAB IV	10.
11.	31 Juli 2023	Konsultasi BAB IV dan BAB V	11.
12.	31 Juli 2023	Konsultasi Revisi dan ACC Keagamaan BAB IV	12.
13.	3 Agustus 2023	Konsultasi Revisi BAB IV dan BAB V	13.
14.	4 Agustus 2023	ACC Seminar Hasil	14.
15.	5 September 2023	Konsultasi Revisi Seminar Hasil	15.



**KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI
MAULANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI**

Jl. Gajayana No.50 Dinoyo Malang Telp. / Fax. (0341)558933

16.	11 September 2023	ACC Sidang Skripsi	16.
17.	02 Oktober 2023	ACC Keagamaan	17.
18.	02 Oktober 2023	ACC Keseluruhan	18.

Malang, 02 Oktober 2023

Mengetahui,

Ketua Program Studi Matematika



Dr. Elly Susanti, M.Sc.

NIP. 19741129 200012 2 005