



**Penggunaan Metode *Multilayer Perceptron* untuk Identifikasi
Sexual harrassment Pada Media Sosial X (Twitter)**

Farhan Wildan Giffary

22917007

Tesis diajukan sebagai syarat untuk meraih gelar Magister Komputer

Konsentrasi Forensika Digital

Program Studi Informatika Program Magister

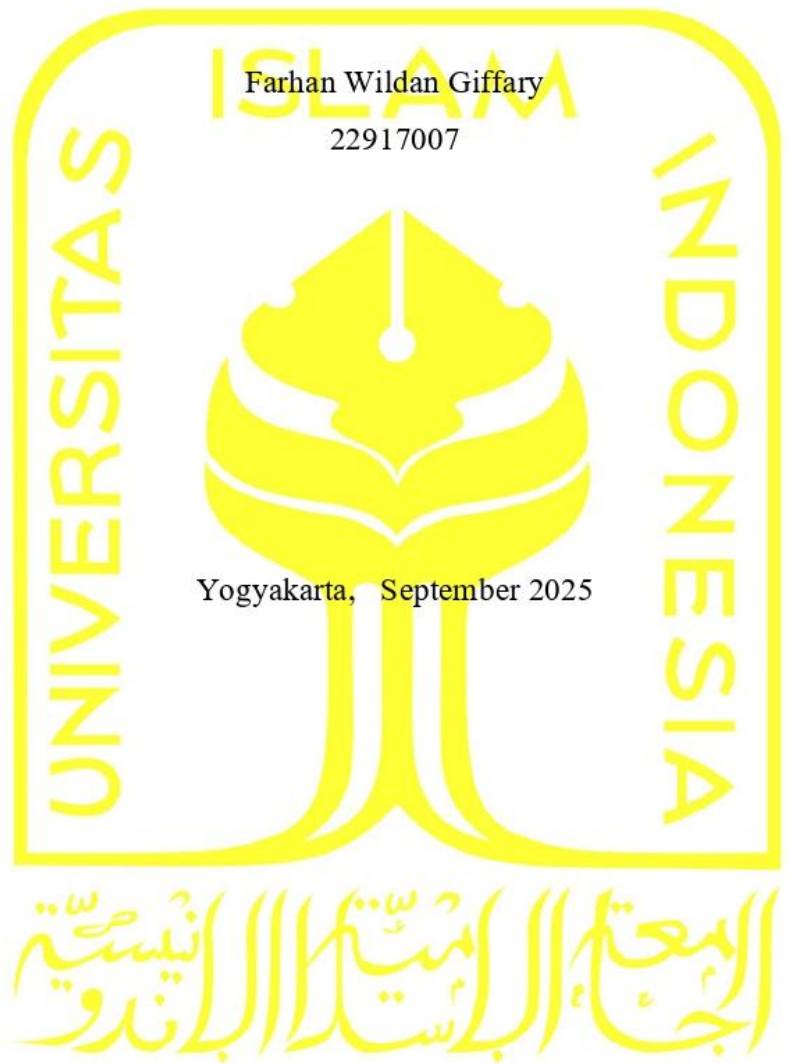
Fakultas Teknologi Industri

Universitas Islam Indonesia

2025

Lembar Pengesahan Pembimbing

**Penggunaan Metode *Multilayer Perceptron* untuk Identifikasi
Sexual harrassment Pada Media Sosial X (Twitter)**



Pembimbing

Prayudi

Dr. Yudi Prayudi, S.Si., M.Kom

Lembar Pengesahan Penguji

**Penggunaan Metode *Multilayer Perceptron* untuk Identifikasi
Sexual harrassment Pada Media Sosial X (Twitter)**

Farhan Wildan Giffary

22917007

Yogyakarta, September 2025

Tim Penguji,

Dr. Yudi Prayudi, S.Si., M.Kom

Ketua



Dr. Ahmad Luthfi, S.Kom., M.Kom

Anggota I



Irving Vita Paputungan, S.T., M.Sc., Ph.D

Anggota II



Mengetahui,

Ketua Program Studi Informatika Program Magister

Fakultas Teknologi Industri

Universitas Islam Indonesia



Irving Vita Paputungan, S.T., M.Sc., Ph.D,

Abstrak

Penggunaan Metode *Multilayer Perceptron* untuk Identifikasi *Sexual harrassment* Pada Media Sosial X (Twitter)

Pelecehan seksual di media sosial merupakan fenomena yang semakin mengkhawatirkan, terutama di platform seperti Twitter yang memungkinkan penyebaran konten secara cepat dan luas. Penelitian ini bertujuan mengembangkan metode deteksi teks pelecehan seksual dengan pendekatan forensik digital berbasis standar NIST SP 800-86. Data dikumpulkan melalui proses *Crawling* yang menghasilkan 1.000 *tweet*, masing-masing berlabel “mengandung pelecehan seksual” atau “tidak mengandung pelecehan seksual”. Tahap *Preprocessing* dilakukan secara berurutan melalui *Cleaning*, normalisasi teks, penghapusan *Stopwords*, dan *Stemming*. Selanjutnya, pembobotan kata dilakukan menggunakan dua metode representasi, yaitu *Term Frequency* (TF) dan *Term Frequency–Inverse Document Frequency* (TF-IDF) berbasis *Unigram*, sebelum masuk ke tahap klasifikasi menggunakan algoritma *Multilayer Perceptron* (MLP). Hasil pengujian menunjukkan bahwa penggunaan fitur TF menghasilkan akurasi sebesar 97,5%, dengan nilai *F1-Score* 0,975, presisi 0,99, dan *recall* 0,961. Sementara itu, penggunaan fitur TF-IDF menghasilkan akurasi sebesar 96,5%, dengan nilai *F1-Score* 0,965, presisi 0,98, dan *recall* 0,951. Meskipun perbedaan kinerja keduanya relatif kecil, hasil ini memperlihatkan bahwa TF murni lebih sesuai dengan karakteristik dataset *tweet* yang memiliki kata kunci spesifik dan berulang, sementara TF-IDF tetap relevan untuk kasus dengan distribusi kata yang lebih bervariasi. Penelitian ini membuktikan bahwa integrasi proses forensik digital dengan teknik *text mining* dan algoritma MLP dapat menjadi solusi potensial untuk mendeteksi pelecehan seksual di media sosial. Kontribusi utama penelitian ini terletak pada penerapannya di bidang forensik digital, khususnya pada tahap analisis bukti digital berbasis teks, yang mampu mendukung investigasi dan penegakan hukum di ranah digital. Dengan akurasi tinggi, model ini diharapkan dapat membantu investigator dalam memilah bukti secara lebih cepat dan akurat, sekaligus memperkuat upaya pencegahan serta perlindungan korban pelecehan seksual daring melalui pemanfaatan teknologi yang efektif dan adaptif.

Kata kunci

Barang bukti, NIST, *Multilayer Perceptron*, X (Twitter), *Digital Investigation*

Pernyataan Keaslian Tulisan

Dengan ini saya menyatakan bahwa tesis ini merupakan tulisan asli dari penulis, dan tidak berisi material yang telah diterbitkan sebelumnya atau tulisan dari penulis lain terkecuali referensi atas material tersebut telah disebutkan dalam tesis. Apabila ada kontribusi dari penulis lain dalam tesis ini, maka penulis lain tersebut secara eksplisit telah disebutkan dalam tesis ini.

Dengan ini saya juga menyatakan bahwa segala kontribusi dari pihak lain terhadap tesis ini, termasuk bantuan analisis statistik, desain survei, analisis data, prosedur teknis yang bersifat signifikan, dan segala bentuk aktivitas penelitian yang dipergunakan atau dilaporkan dalam tesis ini telah secara eksplisit disebutkan dalam tesis ini.

Segala bentuk hak cipta yang terdapat dalam material dokumen tesis ini berada dalam kepemilikan pemilik hak cipta masing-masing. Apabila dibutuhkan, penulis juga telah mendapatkan izin dari pemilik hak cipta untuk menggunakan ulang materialnya dalam tesis ini.



Daftar Publikasi

Kontributor	Jenis Kontribusi
Farhan Wildan Giffary	Mendesain eksperimen (60%) Menulis <i>paper</i> (70%)
Dr. Yudi Prayudi, S.Si., M.Kom	Mendesain eksperimen (40%) Menulis dan mengedit <i>paper</i> (30%)

Halaman Kontribusi

Penelitian ini dapat terselesaikan berkat bimbingan, doa, serta masukan dari berbagai pihak dalam setiap tahapan, mulai dari penyusunan proposal, seminar progres, hingga seminar pendadaran. Dukungan yang diberikan sangat berarti dalam penyempurnaan metodologi, analisis, maupun implementasi penelitian ini.

Ucapan terima kasih secara khusus saya sampaikan kepada:

1. Bapak Dr. Yudi Prayudi, S.Si., M.Kom., selaku dosen pembimbing yang telah memberikan arahan, masukan, dan bimbingan yang penuh kesabaran serta ketelitian.
2. Ibu Erika Ramadhani, S.T., M.Eng. yang telah memberikan masukan berharga dalam proses penelitian, khususnya dalam penyempurnaan metodologi dan analisis data.
3. Bapak Dr. Ahmad Luthfi, S.Kom., M.Kom. yang turut memberikan saran konstruktif dan dorongan moral sehingga penelitian ini dapat terselesaikan dengan baik.
4. Bapak Irving Vitra Paputungan, S.T., M.Sc., Ph.D., yang telah memberikan pandangan kritis dan masukan mendalam yang sangat membantu dalam penyempurnaan kualitas penelitian ini.

Tanpa bimbingan, masukan, dan doa dari Bapak dan Ibu, penelitian ini tidak akan dapat terselesaikan sebagaimana mestinya.

Halaman Persembahan

Bismillahirrahmanirrahim

Puji syukur ke hadirat Allah SWT atas segala limpahan rahmat, karunia, dan hidayah-Nya. Atas izin dan ridho Allah SWT, penulis dapat menyelesaikan tesis ini. Dengan penuh kerendahan hati, karya ini penulis persembahkan kepada:

1. Orang tua tercinta, yang telah menjadi sumber doa, semangat, dan kasih sayang tanpa batas. Segala pengorbanan dan dukungan moril maupun materiil menjadi pilar utama dalam setiap langkah penulis hingga dapat menyelesaikan penelitian ini.
2. Pusat Studi Forensika Digital (PUSFID) Universitas Islam Indonesia, yang telah memberikan wadah pembelajaran, arahan, serta inspirasi dalam mendalami bidang forensik digital.
3. PT Hastama Technologies Indonesia dan Isafe, atas dukungan, kesempatan, serta ruang pengembangan yang sangat berharga bagi penulis dalam mengintegrasikan penelitian ini dengan kebutuhan nyata di dunia industri.

Semoga karya sederhana ini dapat menjadi bentuk balas budi, sekaligus pijakan untuk memberikan manfaat yang lebih luas bagi ilmu pengetahuan, masyarakat, dan dunia kerja.

Kata Pengantar

Bismillahirrahmanirrahim.

Alhamdulillah rabbil ‘alamin, puji syukur penulis panjatkan ke hadirat Allah Subhanahu wa Ta’ala atas segala limpahan rahmat, karunia, dan hidayah-Nya. Berkat izin dan pertolongan-Nya, penulis akhirnya dapat menyelesaikan tesis dengan judul “Deteksi Teks Pelecehan Seksual di Media Sosial dengan Pendekatan Forensik Digital Berbasis *Multilayer Perceptron*”. Tesis ini disusun sebagai salah satu syarat untuk menyelesaikan program Magister di bidang Ilmu Komputer. Proses penyusunan tesis ini tidak terlepas dari berbagai tantangan, baik dalam pengumpulan data, pengolahan, hingga analisis hasil. Namun, berkat doa, dukungan, dan bantuan dari berbagai pihak, penulis dapat menyelesaikan penelitian ini dengan sebaik-baiknya.

Ucapan terima kasih yang tulus penulis sampaikan kepada:

1. Orang tua tercinta, yang senantiasa memberikan cinta, doa, dukungan moral, dan motivasi tanpa henti dalam setiap langkah penulis.
2. Bapak Dr. Yud Prayudi, S.Si., M.Kom, Ibu Erika Ramadhani, S.T., M.Eng., dan Bapak Dr. Ahmad Luthfi, S.Kom., M.Kom., dan Bapak Irving Vitra Paputungan, S.T., M.Sc., Ph.D. selaku dosen pembimbing yang telah dengan sabar memberikan arahan, masukan, dan bimbingan selama proses penelitian ini berlangsung.
3. Rekan-rekan di PUSFID, serta pihak dari PT Hastama Technologies Indonesia dan iSafe, yang turut memberikan kontribusi, dukungan, dan kesempatan dalam melancarkan penelitian ini.

Penulis menyadari bahwa tesis ini masih jauh dari kesempurnaan. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi perbaikan di masa mendatang. Semoga karya ini dapat memberikan manfaat, baik secara akademis maupun praktis, serta berkontribusi bagi perkembangan ilmu pengetahuan khususnya dalam bidang forensik digital dan upaya pencegahan pelecehan seksual di ranah media sosial.

Yogyakarta, Agustus 2025

Penulis

DAFTAR ISI

Lembar Pengesahan Pembimbing	i
Lembar Pengesahan Penguji.....	ii
Abstrak	ii
Pernyataan Keaslian Tulisan	iii
Daftar Publikasi	iv
Halaman Kontribusi.....	v
Halaman Persembahan	vi
Kata Pengantar.....	vii
DAFTAR ISI	iii
DAFTAR TABEL	vi
Glosarium	vii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	2
1.3 Tujuan dan Manfaat Penelitian.....	3
1.4 Manfaat Penelitian.....	3
1.5 Batasan Masalah	3
1.6 Metode Penelitian	3
1.7 Sistematika Penelitian	5
BAB II Tinjauan Pustaka	7
2.1 Literatur Review	7

2.1.1	Penelitian Terdahulu.....	7
2.1.2	Bibliometrik Penelitian Terdahulu	10
2.2	Landasan Teori	12
2.2.1	Forensik Digital	12
2.2.2	<i>Sexual harrassment</i>	13
2.2.3	X (Twitter).....	14
2.2.4	NIST (National Institute of Standards and Technology).....	14
2.2.5	<i>Text Preprocessing</i>	15
2.2.6	Pembobotan <i>Term Frequency</i> -Inverse Document Frecuency (TF-IDF)	16
2.2.7	<i>Multilayer Perceptron</i>	17
BAB III	Metodologi	19
3.1	<i>Collection</i>	20
3.2	Examination.....	21
3.3	Analysis	22
3.4	Reporting	24
BAB IV	Hasil dan Pembahasan	25
4.1	<i>Collection</i>	25
4.2	Examination.....	30
4.2.1	Pre-Proccessing	30
4.2.2	Text Normalization.....	31
4.2.3	Standart Text	32
4.3	Analysis	34
4.3.1	<i>Term Frequency</i> (TF) Feature Extraction.....	35
4.3.2	<i>Term Frequency–Inverse Document Frequency</i> (TF-IDF) Feature Extraction...	36
4.4	Reporting	37
4.4.1	Hasil Pengujian dengan Fitur TF.....	38
4.4.2	Hasil Ekstraksi Fitur <i>Term Frequency – Inverse Document Frequency</i> (TF-IDF)	39
4.4.3	Perbandingan Kinerja Model TF dan TF-IDF	40

4.5	Analisis Dampak dan Kegunaan	42
4.6	Simulasi Tambahan dengan Variasi Keyword	44
BAB V	Kesimpulan dan Saran	47
5.1	Kesimpulan.....	47
5.2	Saran	48

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	8
Tabel 2. 2 Distribusi Publikasi Berdasarkan Tahun	10
Tabel 3. 1 Rincian Data	20
Tabel 4. 1 Ringkasan Perbandingan Kinerja Model	41
Tabel 4. 2 Ringkasan Perbandingan Metrik Kinerja Model	42

DAFTAR GAMBAR

Gambar 1. 1	Alur Metode Penelitian	4
Gambar 2. 1	Jejaring Kolaborasi Penulis.....	11
Gambar 2. 2	Contoh <i>Tweet</i> atau komentar yang mengandung unsur <i>Sexual Harassment</i>	14
Gambar 2. 3	Tahapan <i>Framework</i> NIST	14
Gambar 2. 4	Arsitektur <i>Multilayer Perceptron</i>	17
Gambar 3. 1	Kerangka Penelitian	19
Gambar 3. 3	Tahapan Analisis <i>Text mining</i>	22
Gambar 3. 4	<i>Flowchart</i> alur analisis TF-IDF	23
Gambar 4. 1	Hasil Data <i>Crawling</i>	25
Gambar 4. 2	Labeling Hasil Database	27
Gambar 4. 3	Proses Instalasi <i>Library</i>	28
Gambar 4. 4	Proses Import Library	28
Gambar 4. 5	Proses <i>Read</i> Data	29
Gambar 4. 6	<i>Cleaning Tweet</i> Mentah	30
Gambar 4. 7	Stopword Removal & Normalisasi Kata.....	31
Gambar 4. 8	<i>Stemming</i>	32
Gambar 4. 9	Fungsi Utama Pre-Processing	33
Gambar 4. 10	Contoh hasil <i>Preprocessing</i>	34
Gambar 4. 11	Ekstraksi Fitur <i>Term Frequency</i> (TF)	36
Gambar 4. 12	Ekstraksi Fitur TF-IDFa.....	36
Gambar 4. 13	<i>Confusion Matrix</i> MLP dengan Fitur TF	38
Gambar 4. 14	<i>Confusion Matrix</i> Hasil Klasifikasi Menggunakan Fitur TF-IDF	39
Gambar 4. 15	<i>Confusion Matrix</i> Hasil Klasifikasi Menggunakan TF	40
Gambar 4. 16	<i>Confusion Matrix</i> Hasil Klasifikasi Menggunakan TF-IDF	41
Gambar 4. 17	Metrik Evaluasi Model.....	43
Gambar 4. 18	Alur Proses Pengumpulan dan Analisis Data.....	43
Gambar 4. 19	Contoh Hasil Prediksi pada Data Unlabeled 1	44
Gambar 4. 20	Contoh Hasil Prediksi pada Data Unlabeled 2	45
Gambar 4. 21	Contoh Hasil Prediksi pada Data Unlabeled 3	45
Gambar 4. 22	Contoh Hasil Prediksi pada Data Unlabeled 4.....	45

Glosarium

<i>Sexual Harassment</i>	- Pelecehan seksual, tindakan yang bersifat seksual dan tidak diinginkan oleh korban, yang dapat terjadi dalam bentuk verbal, nonverbal, atau fisik.
<i>Digital Forensics</i>	- Proses pengumpulan, analisis, dan pelestarian bukti digital untuk kepentingan hukum atau investigasi.
NIST SP 800-86	- Standar dari National Institute of Standards and Technology yang memberikan panduan dalam melakukan proses forensik digital.
<i>Crawling</i>	- Proses pengambilan data secara otomatis dari internet menggunakan skrip atau program tertentu.
<i>Preprocessing</i>	- Tahap awal dalam pengolahan teks untuk membersihkan dan menyiapkannya sebelum masuk ke tahap analisis.
<i>Cleaning</i>	- Proses menghapus elemen yang tidak relevan dari teks, seperti URL, tanda baca, angka, dan karakter berulang.
<i>Text Normalization</i>	- Proses mengubah kata tidak baku, singkatan, atau slang menjadi bentuk baku.
<i>Stemming</i>	- Proses mengubah kata ke bentuk dasar untuk menyamakan variasi kata.
<i>Stopwords</i>	- Kata-kata umum yang sering muncul dalam teks namun tidak memberikan kontribusi signifikan pada analisis.
TF (<i>Term Frequency</i>)	- Ukuran yang menghitung seberapa sering sebuah kata muncul dalam dokumen.
TF-IDF (<i>Term Frequency – Inverse Document Frequency</i>)	- Metode pembobotan kata yang mempertimbangkan frekuensi kata dalam dokumen dan kelangkaannya di seluruh korpus.
<i>Unigram</i>	- Model n-gram yang menggunakan satu kata sebagai unit analisis.
MLP (<i>Multilayer Perceptron</i>)	- Model Artificial Neural Network yang terdiri dari beberapa lapisan untuk melakukan klasifikasi atau regresi.
ReLU (<i>Rectified Linear Unit</i>)	- Fungsi aktivasi dalam jaringan saraf yang banyak digunakan karena efisiensinya dalam pelatihan model.
<i>Confusion Matrix</i>	- Tabel yang menunjukkan kinerja model klasifikasi dengan membandingkan prediksi dan label sebenarnya.

Precision

- Ukuran akurasi prediksi positif model.

Recall

- Ukuran kemampuan model untuk menemukan semua data positif yang relevan.

F1-Score

- Rata-rata harmonis dari presisi dan *recall*, digunakan untuk menilai keseimbangan keduanya

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam beberapa tahun terakhir, pengguna media sosial X (Twitter) di telah meningkat. Jumlah pengguna media sosial X (Twitter) meningkat 71,2% hingga 2023 dengan total pengguna 25,25 Juta dengan menempati peringkat 4 terbanyak dunia (Mutia Annur, 2023). Tidak mengherankan bahwa peningkatan ini terjadi, mengingat bagaimana *platform* ini telah mengubah cara orang berkomunikasi, bekerja sama, dan menyebarkan informasi di era digital. Puluhan juta orang secara teratur terlibat dalam berbagai interaksi *online*. Salah satu sumber bukti digital yang paling penting dalam investigasi forensik media sosial, seperti pemeriksaan latar belakang dengan volume data yang sangat besar yang dihasilkan (Bokolo & Liu, 2024). Ini mendorong penelitian tentang metode dan teknologi baru untuk menyelidiki dan mengurangi kejahatan serta perilaku antisosial.

Sexual harrassment menjadi lebih umum seiring dengan kemajuan teknologi dan semakin banyaknya orang yang menggunakan media sosial. Pelecehan seksual merupakan tindakan yang sering terjadi di masyarakat dan menyakiti para korbannya. Seseorang yang pernah mengalami perlakuan pelecehan seksual oleh orang-orang di sekitarnya akan memberikan trauma psikologis dan berdampak negatif pada pembentukan kepribadiannya. Pelecehan seksual dapat berupa konten seksual, membuat lelucon yang menjurus ke arah seksualitas dan penghinaan terhadap bagian tubuh seseorang, serta melakukan kontak fisik berupa sentuhan atau sejenisnya (Budiman, Zaatsiyah, Niswah, Muhanna, & Faizi, 2020). Oleh karena itu, deteksi dini terhadap *sexual harrassment* menjadi penting untuk menghindari konsekuensi yang lebih parah.

Framework dari *National Institute of Standards and Technology* (NIST) merupakan sebuah standar yang umum digunakan untuk manajemen risiko keamanan siber. Kerangka kerja ini menguraikan metode untuk mengidentifikasi, melindungi, mendeteksi, merespons, dan memulihkan diri dari ancaman siber. NIST memiliki aturan tentang cara menangani bukti elektronik. Aturan tersebut merupakan tentang cara menemukan, mengumpulkan, menjaga keamanan, melihat, dan melaporkan barang bukti yang ditemukan. Aturan-aturan ini masih penting untuk menangani bukti digital saat ini (Faizal & Luthfi, 2024). Kerangka

kerja ini dapat digunakan untuk membuat sistem deteksi menyeluruh untuk kasus-kasus *sexual harrassment*.

Dari literatur yang didapat, ada banyak metode dan algoritma yang digunakan untuk analisis *sexual harrassment*, diantaranya pada penelitian yang dilakukan oleh (Daga, Gupta, Vardhan, & Mukherjee, 2020) memprediksi jumlah *retweet* dari *tweet* dengan membandingkan 2 teknik *text Preprocessing*, yaitu *bag of words* (TFIDF) dan word embeddings (Doc2Vec). Hasil *Preprocessing* tersebut akan diuji dengan menggunakan beberapa teknik *machine learning* seperti regresi logistik, *support vector machine* (SVM), random forest, *neural network*, dan *multinomial Naive Bayes*. Selain menggunakan teks, model yang dibuat juga menggunakan tanggal pembuatan *tweet* dan jumlah *like* sebagai fitur tambahan dalam model. Hasil penelitian menunjukkan bahwa kombinasi pre-processing Doc2Vec dan random forest menghasilkan akurasi sebesar 62.67%. Pada penelitian ini peneliti mengusulkan algoritma *Multilayer Perceptron* dan TF-IDF, penggunaan algoritma *Multilayer Perceptron* karena memiliki kemampuan dan performa pembelajaran yang lebih cepat dari algoritma lain (Prasetya Wibawa, Lestar, Bella Putra Utama, Tri Saputra, & Nabila Izdihar, 2020). Algoritma *Multilayer Perceptron* dapat diterapkan pada perangkat lunak ini dengan dilakukan beberapa tahapan, meliputi : pengumpulan data, proses *Preprocessing*, proses pembobotan atau tf-idf, membuat model, dan melakukan train (Sintia, Defit, & Nurcahyo, 2021).

Kombinasi antara *framework* NIST dan metode MLP diharapkan dapat menghasilkan sistem deteksi *Sexual harrassment* yang andal dan responsif. Penelitian ini bertujuan untuk mengembangkan sistem identifikasi awal *sexual harrassment* pada media sosial X (Twitter) dengan mengintegrasikan *framework* NIST dan metode MLP. Diharapkan, sistem ini dapat membantu dalam mendeteksi dan mencegah tindakan *Sexual harrassment* secara efektif.

1.2 Rumusan Masalah

Adapun rumusan masalah yang terdapat dalam penelitian ini yaitu sebagai berikut:

- a. Bagaimana mengimplementasikan metode *Multilayer Perceptron* dalam proses identifikasi *sexual harrassment* pada media sosial X (Twitter)?
- b. Sejauh mana efektivitas metode *Multilayer Perceptron* dalam mengklasifikasikan konten *sexual harrassment* pada media sosial X (Twitter)?

1.3 Tujuan dan Manfaat Penelitian

Tujuan masalah dalam penelitian ini yaitu sebagai berikut:

- a. Meningkatkan akurasi deteksi *sexual harrassment* dengan metode *Multilayer Perceptron* dalam analisis bukti digital dan klasifikasi konten.
- b. Membantu pihak berwenang dalam mengumpulkan bukti digital dan menyelesaikan kasus *sexual harrassment* dengan lebih cepat dan akurat.

1.4 Manfaat Penelitian

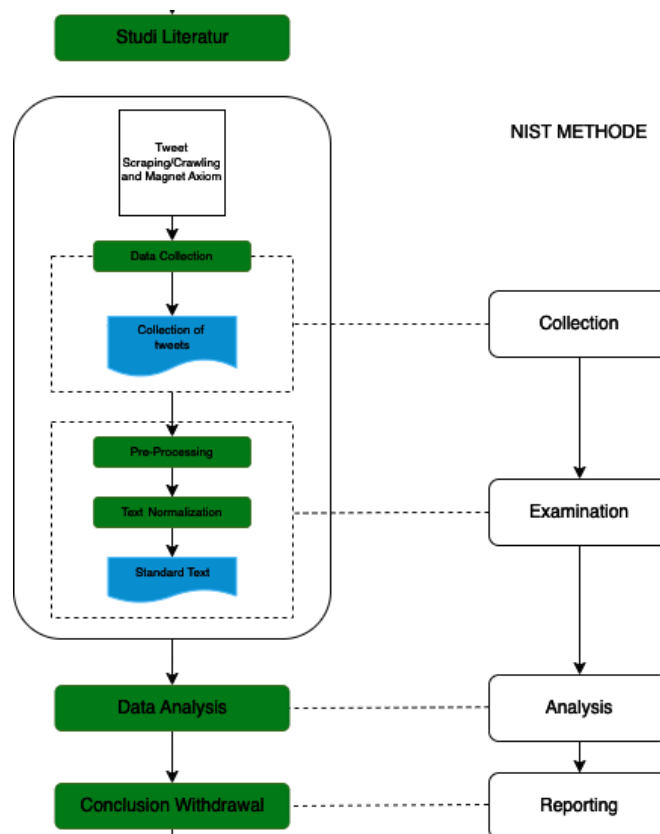
Manfaat dari penelitian ini adalah dapat membantu penyidik untuk mendeteksi kata kata yang mengandung unsur pelecehan seksual pada media sosial, khususnya media sosial X (Twitter) dengan menerapkan algoritma *Multilayer Perceptron*. Penelitian ini juga dapat meningkatkan akurasi pemeriksaan agar lebih efisien dan tertarget yang kemudian akan dijadikan sebagai barang bukti yang sah dalam kasus persidangan.

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini berfokus pada proses analisis dan metode *Multilayer Perceptron* sebagai algoritma klasifikasi untuk mengidentifikasi konten yang mengandung unsur *Sexual harrassment* pada X (Twitter). Adapun aspek lain diluar dari analisis yang dilakukan tidak akan dibahas dalam penelitian ini.

1.6 Metode Penelitian

Pada penelitian ini peneliti menerapkan NIST sebagai kerangka kerja penelitian, adapun beberapa tahapan-tahapan yang dilakukan dalam proses penelitian yaitu sebagai berikut:



Gambar 1. 1 Alur Metode Penelitian

Terdapat beberapa tahapan yang dilakukan dalam penelitian ini, diantaranya diawali dengan studi literatur, *collection*, *examination*, *analysis*, dan *reporting*.

a. Studi Literatur

Dalam penelitian ini, melakukan tahap studi literatur dengan mengeksplorasi berbagai karya penelitian yang berkaitan dengan topik yang sedang diteliti. Tahap ini meliputi pencarian dan pengumpulan teori atau referensi dari sumber-sumber akademis seperti jurnal ilmiah, makalah, buku, dan artikel yang relevan dengan topik penelitian.

b. *Collection*

Dalam penelitian ini, data diperoleh melalui kumpulan data dari X (Twitter) secara global dengan menggunakan kueri tertentu, yaitu *sexual harrassment*. Kata tanya tersebut digunakan untuk mendapatkan data dari sentimen tentang *sexual harrassment*, baik melalui *tweet* biasa di X (Twitter) di Indonesia.

c. *Examination*

Pada tahapan *examination* ini dilakukan pemeriksaan terhadap data yang di dapat pada media sosial X (Twitter) dengan melakukan labelling data dan setelah labelling data selesai dilakukan proses normalisasi teks, yaitu mengubah teks *tweet* menjadi format standar yang lebih mudah dipahami dan dianalisis, seperti menghilangkan tanda baca, mengubah huruf kapital menjadi huruf kecil, dan lain-lain

d. *Analysis/Analisis*

Tahapan analisis dalam penelitian ini diterapkan metode *Multilayer Perceptron*. Metode ini dilakukan untuk mendeteksi teks yang mengandung pelecehan seksual dalam media *social X* (Twitter).

e. *Reporting/Laporan*

Pada tahapan ini dihasilkan laporan yang mencakup ringkasan metode penelitian, hasil, serta kesimpulan dari penelitian yang dilakukan.

1.7 Sistematika Penelitian

Penelitian ini menggunakan sistematika metodologi penelitian atau kerangka penelitian yaitu sebagai berikut:

BAB I: Pendahuluan

Pendahuluan ini menjelaskan latar belakang pentingnya penelitian indentifikasi *sexual harrasment* pada media sosial X (Twitter) dan untuk pemeriksaan forensik digital. Diterangkan pula tujuan penelitian, identifikasi masalah, dan solusi yang diusulkan untuk meningkatkan keamanan di dunia maya.

BAB II: Tinjauan Pustaka

Pada tinjauan pustaka ini mengkaji penelitian sebelumnya terkait identifikasi *sexual harrasment*, metode *Multilayer Perceptron*, dan analisis teks. Selain itu, dibahas konsep dasar tentang forensik digital, bukti digital, teks forensik, dan juga menjelaskan mengenai *Multilayer Perceptron* untuk memperkuat pemahaman terhadap topik ini.

BAB III: Metodologi

Bab ini menjelaskan langkah-langkah penelitian yang digunakan untuk mengidentifikasi *sexual harrasment* menggunakan metode *Multilayer Perceptron*.

Dijelaskan juga teknik analisis teks yang digunakan, seperti *Preprocessing* teks, normalisasi teks, dan TF-IDF.

BAB IV: Hasil Dan Pembahasan

Pada bab ini memaparkan hasil analisis yang dilakukan, menunjukkan hasil yang telah diperoleh, serta mengevaluasi efektivitas dan akurasi metode *Multilayer Perceptron* menggunakan cross validation dan yang diterapkan dalam penelitian ini. Pembahasan hasil mencakup interpretasi temuan dan relevansinya terhadap tujuan penelitian.

BAB V: Kesimpulan Dan Rekomendasi

Bab ini merangkum hasil penelitian dan menjawab rumusan masalah yang telah diidentifikasi. Diberikan pula rekomendasi untuk memperbaiki metode identifikasi *sexual harrasment* dengan metode *Multilayer Perceptron* yang telah diuji, serta saran bagi praktisi forensik digital untuk meningkatkan kesadaran dan kemampuan, guna memastikan integritas bukti digital di masa mendatang.

BAB II

Tinjauan Pustaka

2.1 Literatur Review

2.1.1 Penelitian Terdahulu

Kajian terhadap penelitian terdahulu menjadi langkah penting untuk memahami perkembangan metode, ruang lingkup, dan celah penelitian pada topik yang diangkat. Dalam konteks penelitian ini yang memadukan pendekatan forensik digital dengan klasifikasi teks untuk mendeteksi pelecehan seksual di media sosial terdapat sejumlah studi relevan yang dapat menjadi pijakan.

Sejumlah penelitian telah mengupas deteksi konten bermuatan negatif di media sosial, baik melalui pendekatan machine learning klasik seperti *Naïve Bayes* dan SVM, maupun metode berbasis *deep learning* dan neural network. Di sisi lain, beberapa studi lain mengintegrasikan kerangka kerja forensik digital seperti NIST untuk memperkuat validitas proses pengumpulan bukti digital. Hasil-hasil penelitian ini memberikan gambaran menyeluruh, mulai dari strategi *Preprocessing*, algoritma klasifikasi, hingga penerapan standar forensik yang dapat dipertanggungjawabkan secara hukum.

Selain itu, kajian pustaka ini tidak hanya menyoroti capaian teknis seperti akurasi atau kecepatan pemrosesan, tetapi juga mengidentifikasi kelebihan dan keterbatasan dari masing-masing pendekatan. Dengan demikian, penyusunan penelitian ini tidak sekadar mengulang metode yang sudah ada, melainkan memanfaatkan temuan-temuan sebelumnya sebagai landasan untuk inovasi dan penyempurnaan.

Tabel 2. 1 Penelitian Terdahulu

Sitasi	Metode	Hasil Utama	Kelebihan	Kekurangan
Budiman et al. (2020)	<i>Naïve Bayes</i> , NIST Forensic Acquisition	Mendeteksi sentimen pelecehan seksual di Twitter dengan akurasi cukup tinggi menggunakan dataset berbahasa Indonesia	Mengintegrasikan metode forensik digital NIST dengan klasifikasi teks sehingga hasil dapat dipertanggungjawabkan secara hukum; pembahasan <i>Preprocessing</i> jelas	Tidak menguji metode lain untuk pembandingan; dataset terbatas periode tertentu sehingga variasi slang kurang tercover
Ibrahim et al. (2024)	Neural Network dengan <i>uncertainty estimation</i>	Mendeteksi ujaran kebencian terkait cyberbullying di media sosial secara adaptif	Menggunakan AI modern yang mampu memberikan peringatan saat model tidak yakin; relevan untuk dinamika data media sosial	Fokus pada hate speech, bukan khusus <i>Sexual Harassment</i> ; tidak membahas detail <i>Preprocessing</i>
Ngejane et al. (2021)	Machine Learning + <i>Digital Forensics</i>	Deteksi percakapan predator seksual daring dengan pendekatan forensik	Menggabungkan algoritma ML dengan prosedur forensik; menyoroti kasus predator seksual yang jarang diteliti	Dataset terbatas, sebagian besar bahasa Inggris; model tidak diuji pada bahasa lain
Wakhyudwiono et al.	TF-IDF + <i>Naïve Bayes</i>	Klasifikasi cyber harassment di media sosial	Proses <i>Preprocessing</i> sangat terstruktur; implementasi metode sederhana namun efektif	Tidak memanfaatkan model <i>deep learning</i> yang bisa meningkatkan akurasi
Ardima et al.	SVM & <i>Naïve Bayes</i>	Analisis sentimen komentar media sosial dan	Memberikan perbandingan kuantitatif antar metode; mudah direplikasi	Tidak membahas topik <i>Sexual Harassment</i> ; fokus umum pada sentimen

		perbandingan kinerja algoritma		
Anggraini et al.	NIST Method	Akuisisi bukti digital TikTok berbasis Android	Studi forensik media sosial yang aplikatif dan mengikuti standar resmi	Tidak menggunakan model klasifikasi teks; hanya fokus pada akuisisi bukti
Sintia et al.	Cosine Similarity + TF-IDF	Peningkatan akurasi pengkodean produk dengan pembobotan kata	Memberikan gambaran teknis pembobotan kata yang detail	Tidak relevan langsung pada deteksi harassment; fokus pada e-commerce

Dari Tabel 2.1 terlihat bahwa sebagian penelitian terdahulu cenderung berfokus pada satu aspek, baik pada pengolahan teks semata maupun pada prosedur forensik digital. Penelitian Budiman et al. (2020) dan Ngejane et al. (2021) menjadi pengecualian karena berhasil menggabungkan kedua pendekatan tersebut, meskipun masih terdapat keterbatasan pada lingkup dataset dan metode pembandingan. Hal ini membuka peluang untuk merancang kerangka kerja yang lebih menyeluruh, dengan memperluas cakupan data, meningkatkan teknik *Preprocessing*, serta menguji berbagai model klasifikasi termasuk *Multilayer Perceptron* yang digunakan dalam penelitian ini.

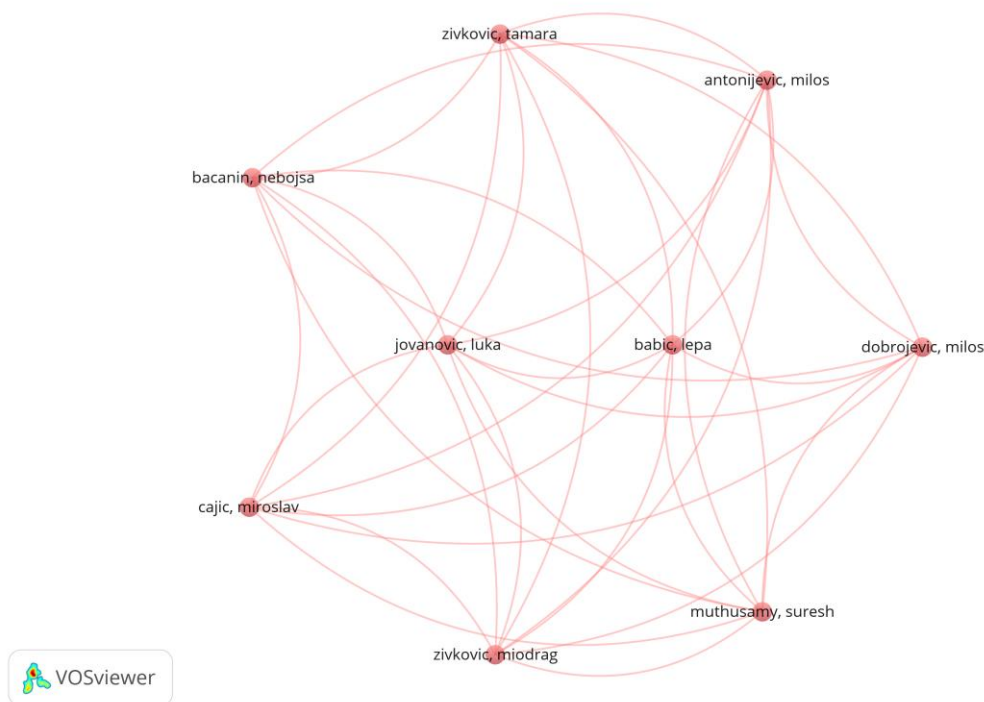
2.1.2 Bibliometrik Penelitian Terdahulu

Analisis bibliometrik dilakukan untuk memahami perkembangan penelitian pada topik *Multilayer Perceptron*, *social media*, *text mining*, dan *Sexual Harassment*. Data diperoleh dari basis data Dimensions.ai melalui serangkaian tahapan pencarian dan penyaringan kata kunci, sehingga hanya publikasi yang relevan dengan ruang lingkup penelitian ini yang dianalisis. Fokus analisis mencakup tren publikasi per tahun, jumlah sitasi yang diperoleh, serta indikator produktivitas dan dampak ilmiah seperti *H-index* dan *G-index*.

Tabel 2. 2 Distribusi Publikasi Berdasarkan Tahun

Tahun	Total Publication	Total Citation	Number of Cited Papers	<i>H-Index</i>	<i>G-Index</i>
2020	2	152	2	2	2
2021	4	87	4	4	4
2022	9	135	6	5	9
2023	5	73	4	3	5
2024	4	20	4	2	4
2025	3	1	1	1	1
TOTAL	27	468	21	17	25

Dari tabel terlihat bahwa produktivitas penelitian mengalami peningkatan signifikan pada tahun 2022 dengan 9 publikasi, disertai jumlah sitasi tertinggi kedua (135 sitasi). Tahun 2020 meskipun hanya menghasilkan 2 publikasi, memiliki jumlah sitasi yang tinggi (152 sitasi), menunjukkan bahwa publikasi di tahun tersebut memiliki pengaruh besar di komunitas ilmiah. Tren ini mengindikasikan adanya peningkatan minat peneliti terhadap topik ini, khususnya setelah 2021, meskipun pada 2024 dan 2025 jumlah sitasi menurun karena faktor kedekatan waktu publikasi yang relatif baru. Nilai *H-index* dan *G-index* menunjukkan kontribusi yang cukup konsisten, dengan *G-index* tertinggi pada 2022.



Gambar 2. 1 Jejaring Kolaborasi Penulis

Visualisasi jejaring kolaborasi penulis yang dihasilkan melalui VOSviewer pada Gambar 2.1 memperlihatkan bahwa seluruh penulis yang teridentifikasi berada dalam satu komponen jaringan tunggal tanpa pembagian cluster yang berbeda. Artinya, meskipun penelitian dilakukan oleh berbagai individu, hubungan kolaborasi antar penulis cukup erat dan saling terhubung. Hal ini bisa mengindikasikan adanya komunitas riset yang solid di bidang ini, namun di sisi lain juga menandakan bahwa potensi kolaborasi lintas kelompok riset atau lintas disiplin masih dapat diperluas.

Keterhubungan yang tinggi ini menjadi indikasi bahwa penelitian pada topik ini masih didominasi oleh lingkaran penulis yang sama atau saling berinteraksi, sehingga peluang untuk memperkenalkan pendekatan dan perspektif baru seperti yang diusulkan dalam penelitian ini terbuka lebar. Dengan mengintegrasikan *text mining* dan machine learning berbasis *Multilayer Perceptron* untuk deteksi *Sexual Harassment* di media sosial, penelitian ini berupaya memberikan kontribusi baru di tengah jaringan penelitian yang sudah ada.

2.2 Landasan Teori

2.2.1 Forensik Digital

Digital forensik merupakan ilmu yang memadukan pengetahuan teknologi komputer dengan metode ilmiah untuk mengungkap dan membuktikan tindak kejahatan di ranah digital. Bidang ini berperan penting dalam proses investigasi modern, di mana setiap tahapan dilakukan secara sistematis guna memastikan integritas dan keabsahan bukti yang ditemukan. Tujuan utamanya adalah memperoleh bukti digital yang sah dan dapat digunakan dalam proses hukum terhadap pihak yang melanggar (Riadi, Yudhana, Caesar, & Putra, 2018). Dalam praktiknya, digital forensik tidak hanya terbatas pada pengumpulan data, tetapi juga mencakup analisis ilmiah terhadap berbagai bentuk kejahatan berbasis teknologi tinggi atau komputer. Analisis ini dilakukan dengan prosedur yang ketat untuk menjamin bukti yang diperoleh benar-benar relevan dan dapat dipertanggungjawabkan secara hukum. Dengan demikian, proses ini memegang peranan strategis dalam memastikan bahwa setiap bukti digital memiliki nilai pembuktian yang kuat di pengadilan (Anggraini, Herman, & Yudhana, 2023).

Forensik digital, atau yang juga dikenal sebagai ilmu forensik digital, merupakan cabang ilmu forensik yang berfokus pada proses pemulihan, pelacakan, dan penyelidikan terhadap data atau barang bukti yang tersimpan dalam perangkat digital. Bidang ini memiliki peran penting dalam mengungkap berbagai bentuk kejahatan digital, mulai dari penipuan daring hingga pelanggaran privasi. Awalnya, istilah ini merujuk secara khusus pada forensik komputer, namun seiring perkembangan teknologi, cakupannya meluas mencakup semua perangkat yang mampu menyimpan data digital mulai dari ponsel pintar, tablet, hingga sistem penyimpanan berbasis awan. Salah satu tujuan fundamental dari investigasi forensik digital adalah memberikan dasar yang kuat untuk mendukung atau membantah suatu hipotesis di hadapan pengadilan, baik dalam perkara pidana maupun perdata. Proses ini menjadi bagian penting dari penemuan bukti elektronik yang sah secara hukum. Melalui prosedur yang sistematis dan terukur, forensik digital memastikan setiap bukti yang dihasilkan memiliki integritas dan reliabilitas yang tinggi, sehingga dapat digunakan sebagai landasan argumen di ranah hukum (Irwansyah & Yudiastuti, 2019).

Secara keseluruhan, digital forensik adalah bidang ilmu yang menggunakan pembuktian hukum dengan menggunakan teknologi dan pengetahuan komputer. Dalam lingkup ini, kemampuan untuk menemukan dan menyelidiki bukti digital yang dapat disimpan dalam berbagai media. Proses ini mencakup analisis perangkat digital untuk mendukung atau menentang hipotesis dalam konteks pengadilan pidana atau perdata. Secara

lebih umum, forensik digital adalah cabang ilmu forensik yang terus berkembang, melibatkan pemulihan dan penyelidikan barang-barang dalam perangkat digital untuk mendukung proses hukum dan investigasi kejahatan komputer.

2.2.2 *Sexual harrassment*

Aktivitas atau perhatian seksual yang tidak diinginkan yang menyebabkan kerugian bagi korban pelecehan disebut pelecehan seksual atau *sexual harrassment*. Pernyataan yang merendahkan tentang orientasi seksual atau seksualitas seseorang, permintaan untuk melakukan tindakan seksual yang diinginkan pelaku, rayuan seksual yang tidak diinginkan, dan komentar atau perilaku yang berkonotasi seksual adalah beberapa bentuk pelecehan seksual. Kegiatan-kegiatan ini dapat dikomunikasikan secara langsung atau secara eksplisit. Korban, bukan pelaku, dianggap sebagai masalah. Karena faktor-faktor ini, korban sering kali tutup mulut untuk waktu yang lama karena mereka takut akan pelecehan lebih lanjut. Mereka tidak membahas masalah ini dengan teman atau kerabat mereka. Proses penyembuhan semakin sulit karena penyangkalan institusi (Triwijati, 2015).

Di sisi lain, pelecehan gender, juga dikenal sebagai pelecehan gender, didefinisikan sebagai perilaku seksual yang kasar, baik secara verbal maupun nonverbal, yang menunjukkan penghinaan, permusuhan, dan merendahkan sikap seseorang tentang gender, identitas gender, atau orientasi seksualnya. Pelecehan gender termasuk gerakan seksual dan cabul, kedipan, dan menampilkan gambar atau objek seksual di tempat kerja. Pelecehan gender juga dapat berupa bahasa, lelucon, atau komentar seksis. Sebagian besar orang setuju bahwa pelecehan seksual termasuk menatap atau melirik seseorang dengan tidak pantas pada bagian tubuh tertentu. Namun, beberapa orang bertindak dengan cara yang berbeda (Sitompul, Sendratari, Sembiring, Arya, & Wirawan, 2023). Pelecehan seksual atau *Sexual Harassment* dapat dikatakan sebagai *Cyberstalking*, seperti penguntit yang dapat mengawasi dan menghubungi korbannya melalui berbagai media digital, adalah salah satu contoh pelecehan *online* lainnya. Menurut Danah Boyd, fitur media sosial mengubah dinamika interaksi seperti itu. Misalnya, kemampuan untuk mengunggah komentar secara anonim pada ide seseorang dapat berdampak pada budaya dan konteks media sosial (Octora, 2019). Salah satu komen atau *tweet* yang mengandung unsur *sexual harrassment* dapat dilihat pada gambar dibawah, dalam konteks *tweet* tersebut sebelumnya terdapat video dari seorang konten creator membuat sebuah video yang isi dari konten tersebut terlihat seorang anak yang sedang makan dan si anak tersebut melihat konten *creator* wanita cantik yang tiba tiba datang dan makan di depan anak tersebut, lalu ada orang yang berkomentar atau *ngetweet*

seolah olah dia mewakili perasaan si anak tersebut.



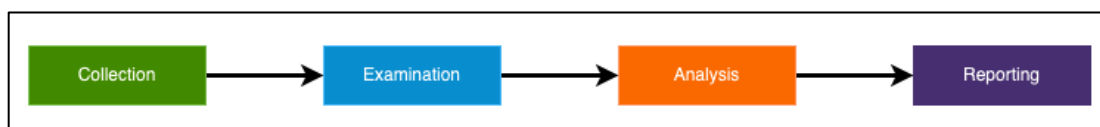
Gambar 2. 2 Contoh *Tweet* atau komentar yang mengandung unsur *Sexual Harassment*

2.2.3 X (Twitter)

X (Twitter) adalah platform media sosial mikroblog yang berfokus pada penyebaran informasi dan interaksi pengguna dalam lingkungan terbuka. *Microblogging* adalah kata yang luas untuk praktik memposting pembaruan status yang sangat singkat. Secara umum, platform *microblogging* seperti X (Twitter) memungkinkan pengguna untuk mengekspresikan pemikiran dan pengalaman mereka dari kehidupan sehari-hari (Hasna & Hanami, 2022). X (Twitter) diklasifikasikan sebagai situs *mikroblog* karena fitur utamanya adalah pesan teks pendek, yang dikenal sebagai *tweet*, dengan panjang maksimum 280 karakter. Fitur *tweet* di Twitter seharusnya memudahkan orang untuk bertukar informasi dengan berfokus pada isi pesan daripada penyampaian yang bertele-tele karena keterbatasan karakter huruf. Twitter mendapatkan popularitas karena fitur *tweet* memungkinkan pengguna untuk membuat komentar cepat dan langsung tentang berbagai topik, termasuk topik pribadi (Budiman et al., 2020).

2.2.4 NIST (National Institute of Standards and Technology)

NIST adalah lembaga federal yang berbasis di Amerika Serikat yang mengembangkan dan mempromosikan standar teknis, rekomendasi, dan pengukuran teknologi. NIST berkontribusi secara signifikan dalam meningkatkan inovasi dan daya saing industri dengan menawarkan standar untuk pertumbuhan keamanan sektor teknologi (Akbar, Falinsa Salsabila, Muhammad Rahdian, & Sidabutar, 2024).



Gambar 2. 3 Tahapan *Framework* NIST

Dari Gambar 2.3 di atas, metode yang dilakukan akan melalui 4 tahapan. Berawal dari *collection*, *examination*, *analysis*, sampai dengan *reporting* (Dasmen, Pratama, Yasir, & Budiman, 2024). Subjek dalam penelitian ini adalah peneliti, adapun objek dari penelitian adalah kasus *sexual harrassment* dengan sumber bukti digital dari media sosial X (Twitter).

2.2.5 Text Preprocessing

Text processing merupakan fondasi awal dalam proses *text mining*, yang berfungsi menyiapkan data teks mentah menjadi bentuk yang terstruktur dan siap diolah lebih lanjut. Dalam konteks *text mining*, proses ini menjadi krusial karena kualitas hasil analisis sangat bergantung pada seberapa baik data awal diproses. Tahap *text processing* melibatkan serangkaian langkah yang bertujuan membersihkan, menyederhanakan, dan memformat teks sehingga dapat dipahami dan diolah oleh algoritma analisis. Dengan kata lain, tahap ini merupakan jembatan yang menghubungkan antara data teks mentah dan proses analisis yang akan dilakukan di tahap berikutnya (A.H. Alamoodi a,*, B.B. Zaidan a,g, A.A. Zaidan a, O.S. Albahri a, K.I. Mohammed a, R.Q. Malik b, E.M. Almahdi a, M.A. Chyad a, Z. Tareq e, A.S. Albahri f, Hamsa Hameed c, 2020). Dalam penelitian ini, tahapan *Preprocessing* yang digunakan meliputi serangkaian prosedur yang dirancang untuk memastikan data bebas dari gangguan seperti simbol yang tidak relevan, kesalahan penulisan, dan format yang tidak seragam. Setiap langkahnya disusun secara berurutan agar hasil akhir benar-benar siap dianalisis dengan metode yang telah direncanakan. Pendekatan ini tidak hanya meningkatkan akurasi model yang digunakan, tetapi juga meminimalkan potensi kesalahan yang mungkin muncul akibat kualitas data yang buruk. Tahapan *Preprocessing* yang digunakan yaitu:

A. Cleansing

Tahap cleansing merupakan tahapan yang dilakukan untuk menghilangkan *noise* atau elemen-elemen yang tidak diperlukan dalam suatu data atau kalimat. Tujuan dilakukannya cleansing adalah untuk meningkatkan dan memperbaiki kualitas data. Tahapan ini dapat mengubah semua huruf dalam dokumen menjadi huruf kecil, hanya huruf “a” sampai dengan “z” yang diterima. Karakter selain huruf dihilangkan dan dianggap delimiter atau pembatas. Selain itu, menghapus tanda baca atau karakter simbol seperti [! “ # \$ % & ‘ () * + , - . / : ; < = > ? @ [] \ ^ _ ` { } | ~], menghapus karakter angka atau digit, dan menghapus spasi yang terdapat kemungkinan suatu kalimat memiliki spasi lebih dari satu (Cahya Kamilla, Priyani, Priskila, & Handrianus Pranatawijaya, 2024).

B. *Slangword*

Slangword atau perbaikan kata tidak baku merupakan proses yang digunakan untuk mengubah kata-kata yang tidak baku menjadi kata baku sesuai dengan Kamus Besar Bahasa Indonesia. Pada perbaikan kata tidak baku terdapat bermacam-macam kata yang dianggap tidak baku, seperti penulisan dengan bahasa modern atau slang (ngga = tidak), penulisan dengan kaidah atau pedoman bahasa (kwalitas = kualitas), penulisan dengan kata singkat (tdk = tidak), dan penulisan dengan kata salah eja (jelsk = jelek). Pada penelitian ini digunakan perbaikan kata tidak baku untuk menangani permasalahan kata tidak baku hanya pada bahasa modern atau slang, penyingkatan kata, dan salah eja saja. Jika terdapat kata tidak baku maka akan diubah menjadi kata baku sesuai dengan kamus. Sehingga didapatkan hasil akhir berupa kalimat dengan menggunakan kata baku.

C. *Stopword*

Proses *stopword removal* bertujuan untuk menghapus kata – kata yang tidak memiliki arti atau tidak mempengaruhi nilai klasifikasi nantinya. Proses *stopword* adalah dengan membuang kata yang diperoleh dari tahap pengecekan pada daftar *stopword*. *Stopword* biasanya berupa kata ganti orang atau kata penghubung. Beberapa contoh dari kata *stopword* seperti ‘aku’, ‘kamu’, ‘yang’, ‘di’, ‘dan’, ‘atau’, ‘tetapi’ dan lain sebagainya. Proses eliminasi *stopword* dalam information retrieval akan membantu dalam mereduksi *feature space* sehingga akan membantu dalam mengurangi waktu dan kompleksitas ruang (Rifaldi, Abdul Fadlil, & Herman, 2023).

D. *Stemming*

Tahap *Stemming* merupakan proses dalam text processing yang digunakan mendapatkan kata dasar dari suatu kata dalam kalimat. Dalam Bahasa Indonesia, kata imbuhan terdiri dari prefiks/awalan, infiks/sisipan dan sufiks/akhiran. Pada penelitian ini proses *Stemming* yang dilakukan hanya akan menangani kata yang memiliki penambahan awalan dan akhiran. Hal ini dilakukan karena proses penambahan infiks dalam bahasa Indonesia jarang terjadi sehingga tidak ada pengaruh yang signifikan terhadap akurasi sistem. Proses *Stemming* dilakukan dengan cara memisahkan masing-masing kata dari kata dasar dan imbuhan, baik awalan maupun akhiran (Albab, Karuniawati P, & Fawaiq, 2023).

2.2.6 Pembobotan *Term Frequency-Inverse Document Frequency (TF-IDF)*

Pembobotan *Term Frequency-Inverse Document Frequency (TF-IDF)* adalah Ekstraksi suatu proses mengubah *term* dari suatu teks menjadi nilai numerik yang dapat dibaca oleh komputer. Salah satu Algoritma yang dapat digunakan adalah *TF-IDF*. Ekstraksi dengan fitur *TF-IDF (Term Frequency and Inverse Document Frequency)* adalah salah satu proses

dari teknik ekstraksi fitur dengan proses memberikan nilai pada masing-masing kata yang ada pada *data training*. Untuk mengetahui seberapa penting sebuah kata mewakili sebuah kalimat, akan diberi nilai perhitungan. Pemberian nilai pada *tf-idf* tergantung besarnya frekuensi kemunculan kata pada dokumen (Antinasari, Setya Perdana, & Fauzi, 2017).

Pada ekstraksi dengan fitur *tf-idf* rumus yang digunakan untuk menghitung nilai (W) dari masing-masing dokumen terhadap kata kunci dengan formula

$$w_{dt} = TF_{dt} * IDF_{ft}$$

Keterangan :

w_{dt} : nilai dokumen ke-d pada kata ke-t

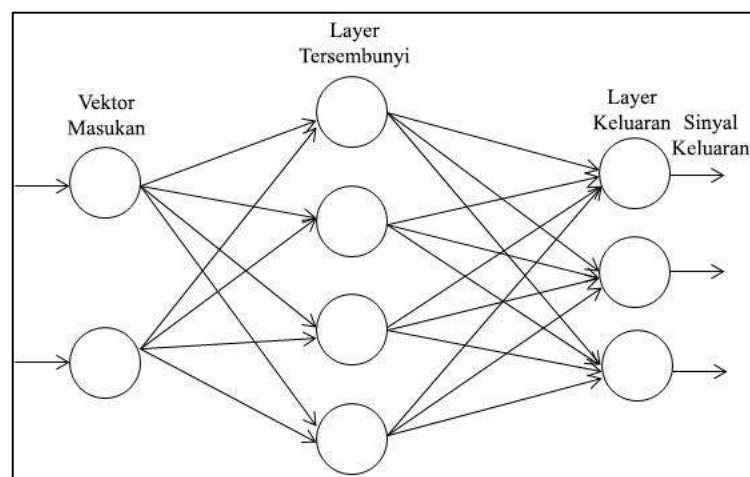
TF_{dt} : jumlah kata yang dicari dalam suatu dokumen

IDF_{ft} : *Inverse Document Frequency* ($\log(\frac{N}{df})$)

N : jumlah dokumen

2.2.7 Multilayer Perceptron

MultiLayer Perceptron (MLP) merupakan turunan dari ANN, yang berupa ANN umpan balik (*feedforward*) dengan salah satu atau lebih layer yang tersembunyi (*hidden layer*). Biasanya, jaringan terdiri atas satu layer masukan, atau setidaknya satu layer *neuron* komputasi di tengah (tersembunyi), dan sebuah layer *neuron* komputasi keluaran. Sinyal masukan ditambahkan dengan arah maju pada layer-per-layer (Ngejane, Eloff, Sefara, & Marivate, 2021). Contoh arsitektur MLP diberikan pada berikut :



Gambar 2. 4 Arsitektur *Multilayer Perceptron*

Pada gambar diatas adalah salah satu layer tersembunyi dengan empat *neuron*, dan satu layer keluaran dengan tiga *neuron*. Sebenarnya ada satu layer lagi di dalam MLP, yaitu

layer masukan berupa vektor masukan. Namun, dalam layer ini tidak ada komputasi, yang dilakukan hanya meneruskan sinyal/vektor masukan yang diterima ke layer di depannya. Setiap layer dalam *Multilayer Perceptron* mempunyai fungsi khusus. Layer tersebut berfungsi menerima sinyal/ vektor masukan dari luar dan mendistribusikannya ke semua *neuron* dalam layer tersembunyi. Layer keluaran menerima sinyal keluaran (atau dengan kata lain, stimulus pola) dari layer tersembunyi dan memunculkan sinyal/nilai/kelas keluaran dari keseluruhan jaringan.

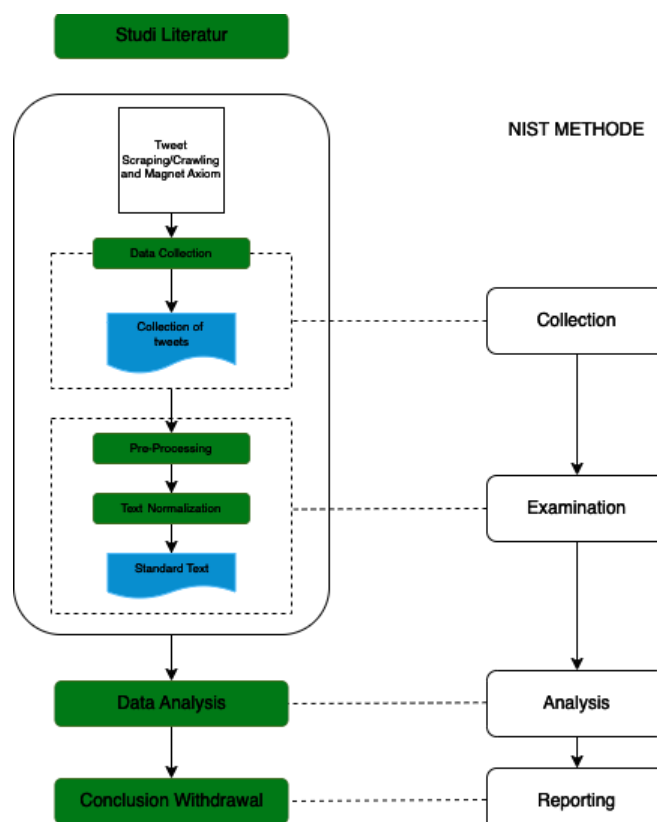
Neuron dalam layer tersembunyi mendeteksi fitur – fitur tersembunyi. Bobot dari *neuron* dalam layer tersembunyi merepresentasikan fitur tersembunyi dalam vektor masukan. Fitur-fitur tersembunyi ini kemudian, digunakan oleh layer keluaran dalam penentuan pola atau kelas keluaran. Dengan satu layer tersembunyi tersebut, kita dapat merepresentasikan sembarang fungsi lanjutan dari sinyal masukan, dan dengan dua layer tersembunyi, fungsi lanjutan dapat direpresentasikan. Layer tersembunyi tersebut menyembunyikan keluaran yang diinginkan. *Neuron* dalam layer tersembunyi tidak dapat diamati melalui perilaku karena masukan/keluaran jaringan secara keseluruhan juga tidak ada cara yang jelas untuk mengetahui keluaran yang diinginkan oleh layer tersembunyi. Dengan kata lain, keluaran yang diinginkan oleh layer tersembunyi ditentukan oleh layer itu sendiri (Palabiyik & Akkan, 2024).

ANN komersial biasanya terdiri dari tiga atau bahkan empat layer, termasuk satu atau dua layer tersembunyi. Setiap layer dapat berisi 10 sampai 1.000 *neuron*. ANN eksperimental dapat mempunyai lima atau bahkan enam layer (tiga atau empat layer tersembunyi) di mana setiap layer menggunakan jutaan *neuron*, tetapi dalam kebanyakan aplikasi menggunakan dua layer (satu layer tersembunyi) karena setiap penambahan satu layer akan meningkatkan beban komputasi secara eksponensial (Ibrahim, Essameldin, & Saad, 2024).

BAB III

Metodologi

Pada penelitian ini melakukan beberapa tahapan yang dilakukan secara sistematis agar menjadi salah satu pedoman dalam menyelesaikan penelitian ini, diantaranya menerapkan NIST (National Institute of Standards and Technology) sebagai kerangka kerja dalam *Digital Forensics*, diawali dengan tahapan *collection*, tahapan *collection* dengan mengambil data dari X (Twitter) menggunakan scraping/*Crawling*, dilanjutkan dengan *examination*, tahapan *examination* melakukan pemeriksaan dan pencarian bukti digital berupa teks dari hasil akuisisi scraping/*Crawling*, dilanjutkan dengan melakukan *analysis*, tahapan *analysis* melakukan analisis dari *evidence* teks yang di dapat untuk mengetahui *sexual harrassment* dengan metode *Multilayer Perceptron*, dan tahapan terakhir yaitu *reporting*, tahapan *reporting* membuat pelaporan hasil dari analisis bukti digital serta melakukan evaluasi bukti digital yang berkaitan dengan kasus *sexual harrassment*.



Gambar 3. 1 Kerangka Penelitian

Dalam tahapan studi literatur, peneliti melakukan eksplorasi dan pemahaman mendalam terhadap karya-karya literatur yang relevan dengan topik penelitian yang sedang dijalankan. Langkah awal melibatkan pencarian literatur dari sumber-sumber akademis, jurnal ilmiah, dan publikasi terkait. Pada tahapan ini memfokuskan perhatian terhadap teori-teori, konsep-konsep, metodologi penelitian, dan temuan-temuan sebelumnya yang dapat memberikan landasan teoritis dan konteks bagi penelitian yang sedang dilakukan.

3.1 Collection

Tahapan ini merupakan bagian dari tahapan identifikasi kebutuhan seperti melakukan pengumpulan data, pengidentifikasi, pelabelan, pencatatan dari sumbernya. Data ini terkait dengan prosedur pemeliharaan integritas data (Umar, Riadi, & Zamroni, 2018). Dalam penelitian ini, data diperoleh melalui kumpulan data dari X (Twitter) secara global dengan menggunakan kueri tertentu, yaitu *sexual harrassment*. Kata tanya tersebut digunakan untuk mendapatkan data dari sentimen tentang *sexual harrassment*, baik melalui *tweet* biasa maupun melalui direct message di X (Twitter) di Indonesia. Selanjutnya mengamati percakapan atau *tweet* teks dari X (Twitter). Pada tahapan ini peneliti menggunakan metode *Crawling* untuk memudahkan dalam pengambil data teks di X (Twitter) karena dapat dilakukan secara otomatis tanpa harus menyalin satu per satu. Data yang sudah didapatkan dari proses *crawling*, selanjutnya diberi label secara manual, sehingga dapat memberikan sebuah kesimpulan label pada suatu kalimat, pada penelitian ini peneliti menggunakan 1000 *tweet* dari X (Twitter) dengan pembagian 500 diberikan label *sexual harrassment* dan 500 label bukan *Sexual Harassment*. Data tersebut kemudian akan digunakan sebagai dataset pada penelitian ini dengan membagi data menjadi 80% sebagai data latih dan 20% sebagai data uji. Rincian jumlah data pada penelitian ini dapat dilihat pada tabel 3.1.

Tabel 3. 1 Rincian Data

No	Label	Jumlah Data
1	<i>Sexual harrassment</i>	500
2	Bukan <i>Sexual harrassment</i>	500
Total		1000

3.2 Examination

Tahap examination pada penelitian ini merupakan langkah dalam proses pemeriksaan, pemeriksaan setelah data digital ditemukan (Nur Faiz, Adi Prabowo, & Fajar Sidiq, 2018). Pemeriksaan merupakan tahap selanjutnya dimana dilakukan pemeriksaan terhadap media sosial X (Twitter) yang akan diekstraksi dan terhadap data hasil *Crawling* dari X (Twitter). Teks percakapan *tweet* dari X (Twitter) kemudian dilakukan proses *text pre-processing* dengan memanfaatkan *cleansing*, *slangword*, *Stopwords* dan *Stemming*. Tahapan pre-processing pada *examination* menjadi penting dimana data yang sudah dikumpulkan akan disiapkan dan dibersihkan agar menjadi teks yang siap untuk di analisis, hal tersebut membuat data yang semula semi-terstruktur atau bahkan tidak terstruktur akan menjadi data yang terstruktur dan siap digunakan. Masing -masing tahapan pre-processing memiliki fungsi dan justifikasi yang saling melengkapi yaitu seperti berikut:

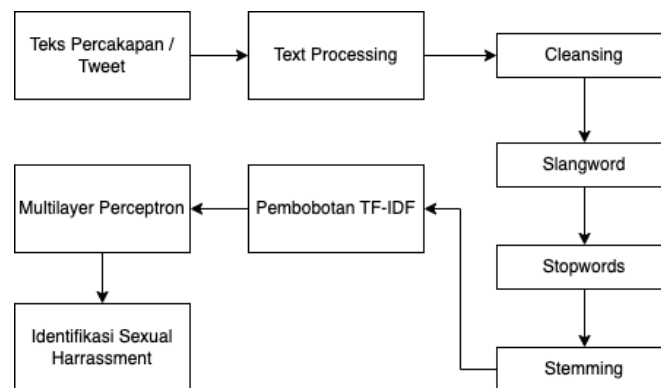
- A. *Cleansing*: merupakan proses mengubah semua huruf pada *tweet* menjadi huruf kecil (*lower case*), menghapus tanda baca, angka, dan *whitespace* untuk menghilangkan *noise* atau elemen – elemen yang tidak diperlukan dalam suatu data atau kalimat.
- B. *Slangword*: merupakan proses mengganti kata yang tidak baku atau singkan dari hasil *cleansing* menjadi kata baku atau aslinya, dalam proses ini akan dilakukan pengecekan per kata dan menggunakan fungsi sinonim dengan mengganti kata yang sebelumnya sudah dibuat daftar kata sinonim dari kata tidak baku atau singkatan yang telah disimpan dalam file csv, jika kata tersebut tidak baku atau singkatan seperti dalam kamus, maka akan diganti dengan sinonim dari kata pada kamus tersebut lalu disimpan dan jika kata tersebut baku atau bukan singkatan, maka kata tersebut akan disimpan.
- C. *Stopwords*: merupakan proses yang bertujuan untuk menghapus kata – kata yang tidak memiliki arti atau tidak berpengaruh terhadap nilai klasifikasi, dengan membuang kata yang diperoleh sehingga nantinya akan menghasilkan kata – kata penting.
- D. *Stemming*: merupakan prses pembentukan kata dasar atau menghilangkan imbuhan dari sebuah kata, proses ini akan mereduksi bentuk guna menghindari ketidakcocokan yang dapat mengurangi nilai *recall*, dimana term – term yang berbeda namun memiliki makna dasar yang sama akan direduksi menjadi satu bentuk (Aldhiansyah, 2020).

Tahapan diatas merupan proses dalam teks *Preprocessing* yang digunakan dalam tahapan selanjutnya yaitu analysis untuk diidentifikasi apakah mengandung unsur *sexual harrasment* atau tidak.

3.3 Analysis

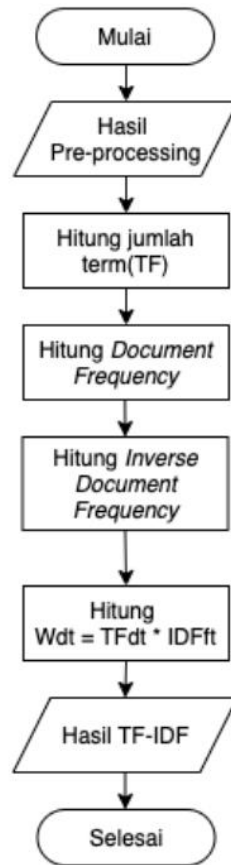
Pada tahap analysis merupakan tahapan melakukan analisis bukti digital yang telah melalui proses pemeriksaan dan kemudian diolah kembali untuk mendapatkan informasi yang berkaitan dengan kasus tersebut (Bintang, Umar, & Yudhana, 2020). Analisis dilakukan untuk mengidentifikasi apakah terdapat bukti-bukti sexual harrasement pada bukti teks percakapan dari direct message Twitter yang telah diekstrak dari proses pemeriksaan sebelumnya.

Tahapan analisis ini dilakukan dengan menggunakan proses *text mining* yang diawali dengan labelling data, kemudian text processing dari proses sebelumnya di examination yang terdiri dari cleansing, slangword, *Stopwords*, dan *Stemming* (Antinasari et al., 2017). Kemudian hasilnya dilakukan pembobotan dengan TF-IDF, dan menerapkan *Multilayer Perceptron* untuk mengidentifikasi indeks *sexual harrasment* pada data percakapan dan *tweet* yang telah didapatkan serta akurasi dari penerapan *Multilayer Perceptron*. Berikut proses analisis yang dilakukan direpresentasikan pada Gambar 3.2.



Gambar 3. 2 Tahapan Analisis *Text mining*

Pada tahapan pembobotan *tf-idf* melakukan proses pemberian nilai atau bobot terhadap setiap term yang ada pada setiap *tweet* yang telah melewati tahapan *pre-processing*, pembobotan ini memiliki tujuan untuk memberikan nilai kepada suatu term. Dimana nilai dari term tersebut akan dijadikan sebagai input pada tahapan klasifikasi, tahapan tersebut dapat dilihat dari gambar sebagai berikut:



Gambar 3. 3 Flowchart alur analisis TF-IDF

Pada gambar 3.3 diatas menjelaskan proses analisis tf-idf dilakukan dengan data teks yang sudah melalui tahap *pre-processing*, kemudian menghitung frekuensi kemunculan setiap kata (term) dalam dokumen tertentu. *Term Frequency* (TF) mengukur seberapa sering sebuah kata muncul dalam satu dokumen. *Inverse Document Frequency* (IDF) mengukur jumlah dokumen yang mengandung term, IDF digunakan untuk mengetahui seberapa umum sebuah kata muncul di seluruh dokumen, semakin jarang sebuah kata muncul di dokumen, semakin tinggi nilai IDF-nya, sehingga kata tersebut dianggap lebih penting.

Setelah proses analisis tf-idf selanjutnya analisis metode *Multilayer Perceptron*, pada tahap analisis ini menggunakan bobot dari tf-idf dimana setiap kata memiliki nilai tinggi, input layer yang digunakan pada penelitian ini terdiri dari 1 dan output layer 1 dengan node pada input layer dan output layer berjumlah sesuai dengan jumlah dataset yang telah di kumpulkan, pemilihan jumlah node pada setiap hidden layer akan dilakukan secara acak, pada analisis MLP juga akan mengetahui hasil akurasi yang di dapat dari proses pengujian model menggunakan cross validation dan *Confusion Matrix*, pengujian dilakukan bertujuan untk mengukur seberapa besar tingkat akurasi yang dihasilkan dari data training dan testing.

3.4 Reporting

Reporting atau pelaporan adalah tahap akhir dari serangkaian kerangka kerja NIST setelah tiga tahap sebelumnya. Hasil analisis bukti digital dan kasus kejahatan yang telah didapatkan untuk selanjutnya dilaporkan pada tahap reporting. Pada tahap ini dilakukan proses pelaporan dan penyerahan hasil dari proses identifikasi teks percakapan dan *tweet* yang diduga mengandung unsur *sexual harrassment* pada aplikasi media sosial Twitter agar hasilnya dapat dijadikan bukti yang valid untuk diproses di ranah hukum, selain itu dilakukan evaluasi terhadap bukti digital apa saja yang didapatkan dan kinerja dari tools yang digunakan, serta pembahasan dari segi tinjauan hukum terkait kasus *sexual harrassment* di media sosial.

BAB IV

Hasil dan Pembahasan

Pada bab ini akan menjelaskan serangkaian hasil dan pembahasan rancangan yang telah dibuat sebelumnya, tahapan ini mengimplementasikan kerangka kerja NIST dengan pembuatan sistem untuk mengidentifikasi *sexual harrassment* menggunakan *Multilayer Perceptron* (mlp), dalam hasil dan pembahasan ini akan diketahui apakah sistem ini telah mencapai tujuan yang diinginkan untuk proses identifikasi dan analisis.

4.1 Collection

Tahap *collection* dalam penelitian ini merepresentasikan fase acquisition pada kerangka kerja NIST, yaitu proses memperoleh bukti digital dari sumber aslinya secara sistematis. Pada konteks penelitian ini, bukti digital berupa data *tweet* yang diperoleh dari platform Twitter melalui proses *Crawling*. Tahap ini sangat penting karena memastikan bahwa data yang diambil benar-benar merefleksikan kondisi lapangan tanpa adanya manipulasi, sehingga dapat diperlakukan sebagai bukti yang sah dalam perspektif forensik digital. Data yang diperoleh kemudian disimpan dalam format terstruktur dengan memperhatikan prinsip integritas bukti, yakni menjaga agar data mentah tidak diubah. Sebagaimana praktik dalam forensik digital, original *evidence* harus tetap terjaga, sementara analisis dilakukan pada *copy evidence*. Oleh karena itu, dataset hasil *Crawling* disimpan dalam bentuk file Excel, yang kemudian digunakan untuk tahap pemeriksaan lebih lanjut.

text
@tanyakanrl Mainin rambut & kenyet jempol kanan tiap mau tidur sampe sekarang lipetan jempol kanan gue agak mengerut wkwkw
@kegblgnunfaedh curiga tuh orang bawa sesuatu yang gak bener
@AndyHuskyyy lengan tunggal di bawahnya
@anxiiousz kenyet di kenyet nyot nyot nyot
@acarnanashh Itu makannya di kenyet ya kak ?
@itsMeErniie Bojong soang Bojong kenyet
Aku akan mulai kerjain ntar sore ø, maaf lama gusy abis dikenyet kenyet sama realita kehidupan

Gambar 4. 1 Hasil Data *Crawling*

Pada Gambar 4.1, data mentah yang terkumpul masih dipenuhi elemen-elemen yang tidak relevan, seperti hashtag, mention, URL, angka, hingga simbol. Dalam perspektif

forensik digital, kondisi ini sepadan dengan bukti digital yang bercampur *noise*, sehingga diperlukan tahap *examination* untuk membersihkannya agar siap dianalisis lebih lanjut. Proses *collection* ini menjadi fondasi penting, sebab kesalahan pada tahap awal berpotensi merusak integritas data dan melemahkan validitas hasil investigasi.

Selain itu, Gambar 4.1 juga memperlihatkan adanya penggunaan istilah atau frasa bermuatan seksual, misalnya kata “kenyot” beserta variasinya. Dalam konteks percakapan digital, istilah ini dapat merujuk pada aktivitas seksual, baik secara eksplisit maupun tersirat. Variasi penggunaannya seperti “di kenyot”, “nyot nyot nyot”, atau “abis dikenyot kenyot” menjadi indikasi adanya gaya bahasa khas pengguna media sosial yang dapat dikategorikan sebagai bentuk pelecehan verbal. Beberapa *tweet* bahkan bersifat multitafsir, sehingga menuntut analisis lebih mendalam untuk membedakan antara penggunaan kata dalam konteks wajar atau yang bermuatan pelecehan seksual.

Proses pelabelan manual kemudian dilakukan terhadap data tersebut. Peneliti meninjau setiap *tweet* dan mengklasifikasikannya ke dalam dua kategori, yaitu *Sexual Harassment* dan *non-Sexual Harassment*. Dari keseluruhan data, digunakan 1.000 *tweet* yang terbagi rata: 500 berlabel *Sexual Harassment* dan 500 lainnya tidak. Tahap pelabelan ini membutuhkan kepekaan linguistik sekaligus pemahaman konteks budaya berbahasa di Indonesia, mengingat bahasa informal, slang, maupun gaya bercanda sering digunakan untuk menyamarkan makna seksual di media sosial.

Dataset yang telah terkumpul dan diberi label kemudian dijadikan dataset utama penelitian. Validitas data ini sangat menentukan kualitas analisis serta akurasi model klasifikasi yang dibangun. Dengan demikian, tahap *collection* tidak sekadar dimaknai sebagai aktivitas teknis mengumpulkan data, melainkan juga sebagai bagian dari proses investigasi awal dalam *Digital Forensics*. Tahap ini berperan menyiapkan bukti digital yang bersih, berintegritas, dan dapat dipertanggungjawabkan pada proses analisis berikutnya.

text	label
@tanyakanrl Mainin rambut & kenyt jempol kanan tiap mau tidur sampe sekarang lipetan jempol kanan gue agak mengkerut wkww	0
@WindiS64491 Buka semua aj nanti aku kenyt atas bawah	1
@anxiousz kenyt di kenyt nyot nyot nyot	0
Kuraba tubuhmu dan ku ciummi mmk mu Disaat kamu tertidur kunikmati tubuhmu Dm buat cewe cs #chatsekskasar #chatseks #moancontent #moancewe #desahancewek #desahanenak #hijabsange #sangeberat #SANGE_AAAAAAAAAAH	1
@AndyHuskyyy lengan tunggal di bawahnya	0
video str3aming s3x ayam kampus ng3nt0t b0k3p terbaru abg c0lm3k muncrat banyak banget	1

Gambar 4. 2 Labeling Hasil Database

Gambar 4.2 memperlihatkan hasil dari proses pelabelan dataset, di mana setiap *tweet* diklasifikasikan secara manual berdasarkan tingkat keterkaitannya dengan unsur *sexual harassment*. Proses ini tidak hanya mengandalkan keberadaan kata-kata tertentu, tetapi juga mempertimbangkan konteks kalimat secara menyeluruh. Beberapa *tweet* tampak menggunakan kata yang ambigu seperti “kenyt”, namun dalam konteks tidak seksual, seperti pada kalimat bercanda atau penggunaan kasual lainnya *tweet* semacam ini diberi label 0. Sebaliknya, ketika kata yang sama digunakan dalam nuansa sugestif atau eksplisit, maka *tweet* tersebut diberi label 1.

Tweet lainnya seperti "Kuraba tubuhmu dan ku ciummi mmk mu..." memperlihatkan penggunaan bahasa yang sangat eksplisit, bahkan dengan upaya penyamaran menggunakan singkatan atau simbol yang lazim digunakan untuk menghindari moderasi algoritma media sosial. *Tweet* seperti ini sangat mencolok karena selain mengandung unsur *sexual harassment* secara verbal, juga berupaya untuk memperluas jangkauan dengan mencantumkan tagar-tagar bertema seksual seperti #chatsekskasar, #moancontent, hingga #sangeberat. Ini menggambarkan praktik nyata bagaimana pelaku menyebarluaskan konten seksual secara terbuka di ruang publik digital.

Melalui proses pelabelan ini, terbentuklah dataset yang lebih representatif terhadap dinamika bahasa di media sosial baik yang eksplisit maupun yang tersirat. Ketelitian pada tahap ini menjadi sangat penting karena pelabelan yang akurat akan menentukan kualitas pembelajaran model klasifikasi, dalam hal ini MLP. Dengan pendekatan yang hati-hati dan kontekstual, tahap labeling memastikan bahwa model tidak hanya belajar dari pola kata

semata, melainkan juga dari nuansa makna yang membedakan antara candaan ringan dengan bentuk pelecahan yang serius. Dalam perspektif *digital forensics*, proses ini sepadan dengan tahap examination, yaitu memilah dan mengklasifikasikan bukti agar lebih siap digunakan pada analisis lanjutan.

```

▼ Install library

[ ] !pip install Sastrawi

Collecting Sastrawi
  Downloading Sastrawi-1.0.1-py2.py3-none-any.whl.metadata (909 bytes)
  Downloading Sastrawi-1.0.1-py2.py3-none-any.whl (209 kB)
    209.7/209.7 kB 5.4 MB/s eta 0:00:00
Installing collected packages: Sastrawi
Successfully installed Sastrawi-1.0.1
```

Gambar 4. 3 Proses Instalasi *Library*

Pada tahap awal, dilakukan instalasi pustaka (*library*) yang dibutuhkan untuk mendukung seluruh proses analisis. Salah satu pustaka penting yang digunakan adalah Sastrawi, yang berfungsi sebagai alat *stemming* untuk bahasa Indonesia. Instalasi ini memastikan bahwa seluruh modul eksternal yang diperlukan dapat dijalankan tanpa kendala di lingkungan pemrograman, sehingga meminimalkan potensi error saat pemanggilan fungsi di tahap berikutnya.

```

▼ Import library

[ ] import warnings
warnings.filterwarnings('ignore')

import pandas as pd
import numpy as np
import re
from collections import Counter
import pickle
from matplotlib import pyplot as plt

import nltk
nltk.download('punkt')
nltk.download('punkt_tab')
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
from sklearn.feature_extraction.text import CountVectorizer, TfidfVectorizer
from sklearn.model_selection import GridSearchCV

from sklearn.neural_network import MLPClassifier
from sklearn.model_selection import train_test_split
from sklearn.model_selection import cross_validate
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score, confusion_matrix
from sklearn.metrics import ConfusionMatrixDisplay

pd.set_option('display.max_columns', 1000)
pd.set_option('display.max_rows', 1000)

factory = StemmerFactory()
stemmer = factory.create_stemmer()

[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Unzipping tokenizers/punkt.zip.
[nltk_data] Downloading package punkt_tab to /root/nltk_data...
[nltk_data] Unzipping tokenizers/punkt_tab.zip.
```

Gambar 4. 4 Proses Import Library

Pada tahap ini, seluruh pustaka (*library*) yang diperlukan untuk penelitian diimpor ke dalam lingkungan kerja. Pustaka tersebut mencakup *Pandas* dan *NumPy* untuk pengolahan data, *Matplotlib* untuk visualisasi, *Sastrawi* untuk proses *stemming* bahasa Indonesia, serta *Scikit-learn* untuk berbagai kebutuhan seperti feature extraction, pemodelan menggunakan machine learning, dan evaluasi model. Selain itu, pustaka *NLTK* digunakan untuk proses tokenisasi teks. Proses import ini memastikan bahwa semua fungsi dan alat yang dibutuhkan tersedia, sehingga setiap tahap analisis dapat berjalan tanpa hambatan teknis.

```
Read data

[ ] df_data = pd.read_excel('XTweetTest.xlsx')
    print(df_data.shape[0])
    df_data.head()
```

	text	label
0	@tanyakani Mainin rambut & kenot jempol ...	0
1	@kegblgrunfaedh curiga tuh orang bawa sesuatu ...	0
2	@Andykuskyyy kengan tunggal di bawahnya	0
3	@anxiousz kenot di kenot nyot nyot nyot	0
4	@acamanashh itu makannya di kenot ya kak ?	0

```
[ ] df_data.label.value_counts(dropna=False)
```

label	count
0	500
1	500

dtype: int64

Gambar 4. 5 Proses *Read Data*

Tahap berikutnya adalah pembacaan dataset menggunakan pustaka *Pandas*, di mana data penelitian diimpor dari file *XTweetTest.xlsx*. File ini berisi kumpulan teks hasil *crawling* beserta label klasifikasinya. Pada gambar ditampilkan struktur awal dataset yang menunjukkan lima entri pertama, termasuk contoh *tweet* yang akan dianalisis.

Selain itu, dilakukan pula pengecekan distribusi label untuk memastikan keseimbangan kelas. Dataset ini terdiri dari 500 data berlabel 0 (tidak mengandung pelecehan seksual) dan 500 data berlabel 1 (mengandung pelecehan seksual). Distribusi yang seimbang ini sangat penting dalam machine learning, karena mencegah model cenderung bias terhadap salah satu kelas. Dalam perspektif forensik digital, keseimbangan data ini dapat dianalogikan sebagai upaya menjaga proporsionalitas bukti, sehingga analisis yang dihasilkan tidak berat sebelah dan tetap memiliki validitas yang kuat ketika dijadikan dasar interpretasi lebih lanjut.

4.2 Examination

Pada tahap ini dilakukan proses pemeriksaan dan penyiapan data hasil *Crawling* agar layak untuk dianalisis lebih lanjut. Pemeriksaan meliputi proses transformasi teks mentah menjadi data yang bersih dan terstruktur melalui tiga tahapan penting, yaitu: *preprocessing*, *text normalization*, dan *standard text transformation*. Ketiga tahap ini berperan penting untuk memastikan bahwa data terbebas dari gangguan *noise* serta siap menjadi input dalam sistem klasifikasi *multilayer perceptron*.

4.2.1 Pre-Processing

Preprocessing merupakan langkah awal yang sangat penting dalam menyiapkan data teks untuk dianalisis lebih lanjut. Data mentah hasil *crawling* dari media sosial X (Twitter) pada dasarnya masih bercampur dengan elemen-elemen yang tidak relevan, seperti tautan URL, mention pengguna, hashtag, angka, hingga tanda baca yang berlebihan. Tidak jarang pula ditemui teks dengan huruf kapital penuh atau pengulangan karakter tidak wajar (misalnya: “haiiiiiiii”). Apabila kondisi ini dibiarkan, maka hasil analisis akan terganggu, dan proses investigasi digital dapat menghasilkan temuan yang keliru.

Untuk itulah digunakan fungsi *Cleaning()*, yang bertugas menormalisasi teks agar siap diekstraksi lebih lanjut. Fungsi ini bekerja dengan mengubah semua huruf menjadi lowercase, menghapus username, URL, simbol, angka, serta karakter berulang. Tujuan dari tahap ini bukan hanya merapikan data, tetapi juga menyusun teks ke dalam bentuk yang lebih terstandar sehingga memudahkan proses ekstraksi data di platform analisis. Dengan data yang sudah rapi, sistem klasifikasi dapat mengenali pola bahasa secara lebih konsisten, sementara peneliti dapat menelusuri bukti digital dengan akurasi lebih tinggi.

```
def cleaning(tweet):
    # lowercase
    normal_tw = tweet.lower()
    # hapus URL
    normal_tw = re.sub('((www\.[^\s]*)|(https?:\/\/[^\s]*))|(pic\.twitter\.com/[^\s]*)', '', normal_tw)
    # hapus @username
    normal_tw = re.sub(r'@[^\s]+', '', normal_tw)
    # hapus hashtag
    normal_tw = re.sub(r'#[^\s]+', '', normal_tw)
    # hapus tanda baca
    normal_tw = re.sub(r'[\W\s]', '', normal_tw)
    # hapus angka
    normal_tw = re.sub(r'\d+', '', normal_tw)
    # remove spasi berlebih
    normal_tw = re.sub(r'\s+', ' ', normal_tw)
    # trim depan belakang
    normal_tw = normal_tw.strip()
    # regex huruf yang berulang kaya haiiii (untuk fitur unigram)
    normal_regex = re.compile(r'(\w)\1{1,}')
    # buang huruf yang berulang
    normal_tw = normal_regex.sub(r'\1\1', normal_tw)
    return normal_tw
```

Gambar 4. 6 *Cleaning Tweet Mentah*

Seperti terlihat pada Gambar 4.6, teks mentah yang semula penuh *noise* dapat dipulihkan menjadi data yang lebih bersih dan homogen. Tahap ini menjadi kunci penting dalam rantai investigasi digital, karena memastikan setiap potongan teks dapat diproses secara sistematis tanpa kehilangan esensinya. Dengan begitu, dataset siap digunakan pada tahap selanjutnya, yaitu normalisasi kata dan transformasi standar, yang akan memperkuat kualitas ekstraksi data untuk analisis forensik maupun klasifikasi berbasis machine learning.

4.2.2 Text Normalization

Setelah tahap *preprocessing* dilakukan, langkah berikutnya adalah normalisasi teks. Tahap ini penting untuk memastikan bahwa teks yang sudah dibersihkan dapat diseragamkan ke dalam bentuk standar, sehingga memudahkan proses ekstraksi data dan analisis lebih lanjut. Dalam praktik investigasi digital, normalisasi berfungsi layaknya proses penyusunan bukti setiap elemen data diberi bentuk yang konsisten agar tidak menimbulkan interpretasi ganda di tahap berikutnya.

Pada penelitian ini, proses normalisasi dilakukan dengan menghapus *stopwords* serta mengganti kata-kata tidak baku, singkatan, dan bahasa slang menjadi bentuk formal. Misalnya, kata “ngga” dinormalisasi menjadi “tidak”, atau kata “gusy” menjadi “gusi”. Proses ini menggunakan kamus berbasis file CSV, yang terdiri dari daftar singkatan, kamus alay, dan daftar *stopwords*. Setiap token dalam *tweet* dicek terhadap kamus tersebut, lalu diubah ke bentuk formal yang sesuai

```
def remove_stopwords_and_normalize(tweet):
    stopwords = pd.read_csv('stopwordsID.csv', header=None)

    df_kamus_singkatan = pd.read_csv('kamus_singkatan.csv')
    df_kamus_alay = pd.read_csv('colloquial-indonesian-lexicon.csv')

    token = nltk.word_tokenize(tweet)
    token_new = []
    for k in token:
        if k in df_kamus_singkatan['singkatan'].values:
            k = df_kamus_singkatan.loc[df_kamus_singkatan['singkatan']
                                       == k, 'asli'].values[0]

        if k in df_kamus_alay['slang'].values:
            k = df_kamus_alay.loc[df_kamus_alay['slang']
                                  == k, 'formal'].values[0]

        if k not in stopwords[0].values:
            token_new.append(k)

    str_clean = ' '.join(token_new)
    return str_clean
```

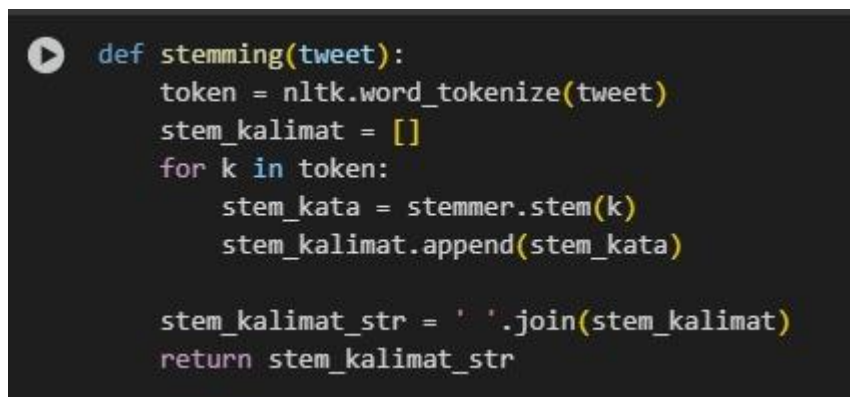
Gambar 4. 7 Stopword Removal & Normalisasi Kata

Sebagaimana ditunjukkan pada Gambar 4.7, hasil normalisasi menghasilkan teks yang lebih konsisten dan formal dibandingkan bentuk mentahnya. Proses ini membantu menyaring kata-kata yang tidak memberikan makna signifikan dalam analisis, sekaligus menegaskan makna kata yang sebelumnya ambigu. Bagi kebutuhan forensik digital, hal ini sangat krusial, karena teks dengan bentuk standar akan lebih mudah diekstraksi, ditelusuri, dan dibandingkan lintas dokumen. Dengan demikian, tahap ini bukan hanya menyederhanakan data, tetapi juga mengoptimalkan kualitas bukti digital yang siap digunakan dalam sistem klasifikasi berbasis *multilayer perceptron*.

4.2.3 Standart Text

Tahap akhir dalam proses Examination adalah transformasi teks ke bentuk standar menggunakan teknik *stemming*. Tujuan dari tahap ini adalah mengembalikan setiap kata ke bentuk dasarnya (*root word*) sehingga berbagai bentuk turunan kata seperti "dimakan", "makanan", atau "pemakan" akan direduksi menjadi satu kata dasar: "makan". Ini penting karena dalam klasifikasi berbasis teks, bentuk kata yang berbeda namun bermakna sama dapat menimbulkan kebingungan bagi model, dan mengakibatkan redundansi fitur.

Dalam penelitian ini, fungsi *Stemming()* memanfaatkan pustaka Sastrawi yang dirancang khusus untuk bahasa Indonesia. Setiap *tweet* yang telah dinormalisasi diproses kata per kata, lalu dikembalikan ke dalam bentuk string utuh. Hasilnya adalah teks dengan kata-kata dalam bentuk dasar yang konsisten, sehingga siap digunakan dalam tahap ekstraksi fitur.



```
def stemming(tweet):
    token = nltk.word_tokenize(tweet)
    stem_kalimat = []
    for k in token:
        stem_kata = stemmer.stem(k)
        stem_kalimat.append(stem_kata)

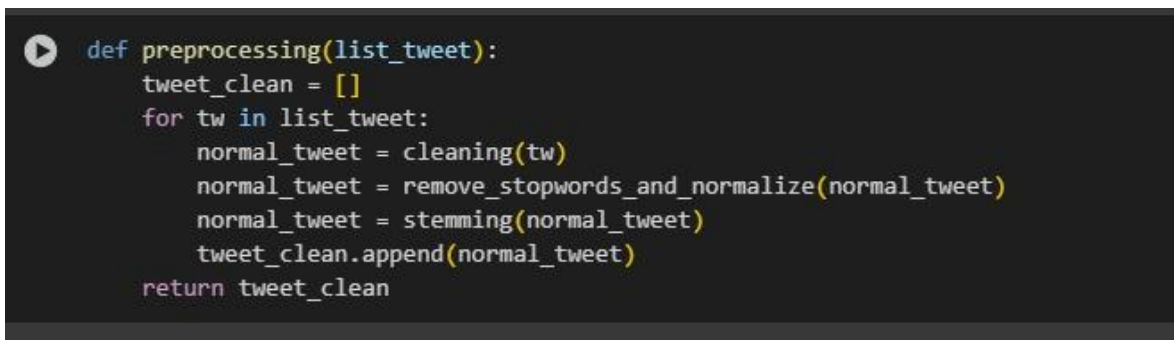
    stem_kalimat_str = ' '.join(stem_kalimat)
    return stem_kalimat_str
```

Gambar 4. 8 *Stemming*

Pada Gambar 4.8 terlihat bagaimana fungsi *Stemming()* menerima input berupa daftar *tweet* yang telah melalui tahap pembersihan dan normalisasi, lalu mengembalikannya

dalam bentuk akhir yang hanya terdiri dari kata-kata dasar. Hasil ini memberikan keuntungan nyata: data menjadi lebih efisien, lebih ringkas, dan semantik yang terkandung di dalamnya semakin kuat. Dengan dihilangkannya imbuhan dan variasi morfologis, model tidak lagi dibebani untuk mempelajari bentuk kata yang berbeda-beda padahal merujuk pada konsep yang sama. Dampaknya, proses klasifikasi dapat berlangsung dengan lebih fokus dan akurat, khususnya dalam mendeteksi konten yang mengandung indikasi *sexual harassment*.

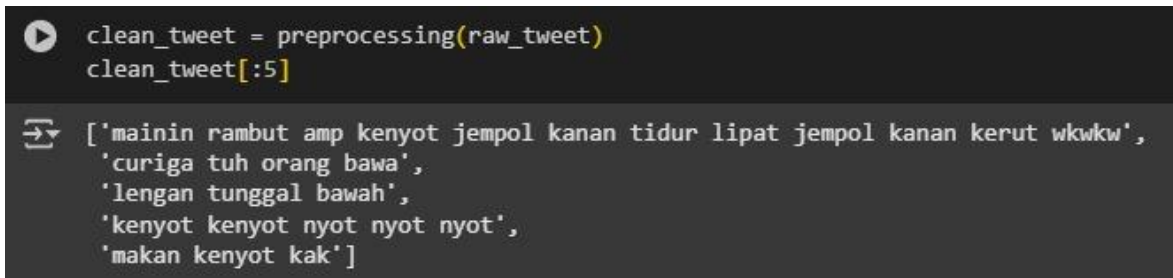
Sebagai penutup dari tahapan examination, seluruh proses mulai dari *cleaning*, normalisasi, hingga *stemming* tidak dapat dipandang sebagai tahapan yang berdiri sendiri, melainkan sebagai bagian dari satu rangkaian yang saling terhubung. Integrasi tersebut dirangkum dalam sebuah fungsi utama bernama *Preprocessing()*, yang secara otomatis mengatur urutan kerja dari awal hingga akhir: membersihkan teks dari *noise*, menghapus unsur yang tidak bermakna, menormalisasi kata tidak baku, dan akhirnya mereduksi kata ke bentuk dasarnya.

A screenshot of a code editor showing a Python function named `preprocessing`. The function takes `list_tweet` as an argument and returns a list of cleaned tweets. The code includes comments for each step: `cleaning`, `remove_stopwords_and_normalize`, and `stemming`.

```
def preprocessing(list_tweet):  
    tweet_clean = []  
    for tw in list_tweet:  
        normal_tweet = cleaning(tw)  
        normal_tweet = remove_stopwords_and_normalize(normal_tweet)  
        normal_tweet = stemming(normal_tweet)  
        tweet_clean.append(normal_tweet)  
    return tweet_clean
```

Gambar 4. 9 Fungsi Utama Pre-Processing

Seperti ditunjukkan pada Gambar 4.9, fungsi *Preprocessing()* dirancang untuk memproses setiap *tweet* secara berkesinambungan dalam bentuk daftar (list). Kode yang dibangun menggabungkan semua tahapan pemeriksaan sebelumnya dan memastikan setiap entri diperlakukan dengan standar yang konsisten. Desain modular ini menjadikan proses lebih efisien, terstruktur, dan dapat dengan mudah diterapkan pada dataset lain dalam skala besar tanpa perlu mengubah logika dasar. Dalam konteks investigasi digital, fungsi semacam ini dapat dipandang sebagai mekanisme otomasi penyusunan bukti, yang memudahkan peneliti maupun praktisi ketika mengekstraksi data dari berbagai platform untuk tujuan analisis maupun pembuktian.



```
clean_tweet = preprocessing(raw_tweet)
clean_tweet[:5]
```

```
['mainin rambut amp kenyot jempol kanan tidur lipat jempol kanan kerut wkwkw',
 'curiga tuh orang bawa',
 'lengan tunggal bawah',
 'kenyot kenyot nyot nyot nyot',
 'makan kenyot kak']
```

Gambar 4. 10 Contoh hasil *Preprocessing*

Gambar 4.10 memperlihatkan hasil nyata dari pelaksanaan fungsi *Preprocessing()* pada lima *tweet* pertama dalam dataset. Teks yang sebelumnya penuh dengan tanda baca tidak perlu, kata-kata tidak baku, simbol, serta struktur kalimat yang tidak konsisten, kini telah berubah menjadi susunan kata yang lebih bersih, ringkas, dan seragam. Perubahan ini tidak hanya bersifat kosmetik, melainkan bersifat fundamental: teks yang sudah terstandarisasi memiliki daya representasi yang jauh lebih kuat ketika digunakan pada tahap analisis. Model klasifikasi tidak terbebani oleh keragaman bentuk kata yang semata-mata variasi penulisan, melainkan dapat berfokus pada makna inti dari pesan yang dianalisis. Hasil ini menegaskan bahwa *preprocessing* bukanlah sekadar langkah teknis awal, melainkan pondasi utama yang memastikan kualitas data tetap terjaga. Peran *preprocessing* menjadi sangat menentukan untuk menyiapkan bukti digital agar lebih siap diekstraksi, ditelusuri, dan diklasifikasikan, sehingga mendukung keberhasilan deteksi otomatis terhadap konten *Sexual Harassment* di ranah media sosial.

4.3 Analysis

Tahap analisis merupakan bagian krusial dalam penelitian ini, karena di sinilah data hasil *preprocessing* yang telah melalui proses pembersihan, normalisasi, dan *stemming* ditransformasikan menjadi bentuk numerik agar dapat dipahami oleh algoritma pembelajaran mesin. Dengan kata lain, teks yang semula berupa rangkaian kata diolah menjadi representasi vektor yang dapat diekstraksi, dipelajari, dan diproses lebih lanjut dalam kerangka klasifikasi.

Pada penelitian ini digunakan dua pendekatan utama untuk melakukan representasi, yaitu *Term Frequency* (TF) dan *Term Frequency–Inverse Document Frequency* (TF-IDF). TF bekerja dengan menghitung frekuensi kemunculan kata dalam sebuah dokumen, sehingga memberikan gambaran pola kata yang sering digunakan. Sementara itu, TF-IDF menambahkan dimensi lain berupa bobot, di mana kata yang terlalu umum di banyak

dokumen diberikan bobot rendah, sedangkan kata yang lebih jarang namun spesifik mendapatkan bobot lebih tinggi.

Proses ekstraksi fitur ini sangat penting dalam konteks investigasi digital. Tanpa bobot yang tepat, model berisiko “tertipu” oleh kata-kata umum seperti “aku” atau “kamu” yang kerap muncul di hampir semua *tweet*, tetapi tidak relevan dalam membedakan apakah teks tersebut mengandung pelecehan seksual atau tidak. Sebaliknya, kata-kata bernuansa seksual seperti “kenyot” atau “ciummi” justru perlu ditangkap dan diberi bobot yang lebih besar, karena sering kali menjadi indikator utama dalam kasus *Sexual Harassment*. Dengan demikian, tahap analisis ini berfungsi sebagai mekanisme penyaringan dan penekanan, sehingga hanya pola yang benar-benar penting yang dipelajari oleh algoritma *multilayer perceptron*.

4.3.1 Term Frequency (TF) Feature Extraction

Proses ekstraksi fitur berbasis *Term Frequency* dilakukan menggunakan pendekatan *CountVectorizer* dari pustaka *Scikit-learn*. Pada tahap ini, teks yang telah melewati *Preprocessing* diubah menjadi matriks numerik, di mana setiap kolom mewakili kata unik dan setiap baris mewakili dokumen (*tweet*). Nilai pada matriks menunjukkan jumlah kemunculan kata dalam dokumen tertentu. Dengan cara ini, data teks yang semula tidak terstruktur dapat ditransformasikan menjadi bentuk yang siap dipelajari oleh algoritma klasifikasi.

Penelitian ini menggunakan parameter *ngram_range=(1,1)* atau *Unigram*, sehingga unit analisis yang digunakan hanyalah kata tunggal. Pendekatan ini dipilih karena lebih sesuai dengan sifat data *tweet* yang relatif pendek dan sering kali mengandalkan kata kunci sebagai penanda makna. Meskipun sederhana, representasi *Unigram* terbukti mampu menangkap pola bahasa yang relevan, terutama ketika berhadapan dengan istilah-istilah bermuatan seksual yang sering muncul secara eksplisit maupun tersirat.

```
...
Function to extract TF (1-gram) features
...

def tf_extraction(text, ngram_start, ngram_end):
    ngram = CountVectorizer(ngram_range=(ngram_start, ngram_end))
    ngram_matrix = ngram.fit_transform(np.array(text)).todense()
    feature_names = ngram.get_feature_names_out()
    return ngram_matrix, feature_names

# unigram features
ngram_feat, feature_names = tf_extraction(clean_tweet, 1, 1)
print(ngram_feat[:3])
print(feature_names[:3])

[[0 0 0 ... 0 0 0]
 [0 0 0 ... 0 0 0]
 [0 0 0 ... 0 0 0]]
['aa' 'abai' 'abal']
```

Gambar 4. 11 Ekstraksi Fitur *Term Frequency* (TF)

4.3.2 *Term Frequency–Inverse Document Frequency* (TF-IDF) Feature Extraction

Setelah proses TF selesai, tahap selanjutnya adalah TF-IDF weighting menggunakan *TfidfVectorizer*. Metode ini tidak hanya menghitung frekuensi kemunculan kata, tetapi juga memberikan pembobotan berdasarkan kelangkaan kata dalam keseluruhan dokumen. Dengan cara ini, kata-kata umum seperti “dan”, “atau”, atau “yang” akan memperoleh bobot rendah, sementara kata-kata yang jarang muncul namun relevan dengan konteks akan memiliki bobot tinggi.

```
[16] ...
Function to extract TF-IDF (1-gram) features
...

def tf_idf_extraction(text):
    vectorizer = TfidfVectorizer()
    tfidf_matrix = vectorizer.fit_transform(np.array(text)).todense()
    return tfidf_matrix

# tf-idf features
tfidf_feat = tf_idf_extraction(clean_tweet)
print(tfidf_feat[:3])

[[0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
```

Gambar 4. 12 Ekstraksi Fitur TF-IDFa

Implementasi pada Gambar 4.12, penelitian ini tetap menggunakan *Unigram* sebagai satuan analisis. Seluruh teks telah dikonversi ke huruf kecil (*lowercase*) pada tahap *Preprocessing*, sehingga parameter *lowercase* pada *TfidfVectorizer* diset ke nilai *False* untuk menghindari redundansi. Hasil TF-IDF berupa matriks numerik padat (*dense matrix*) yang siap digunakan dalam proses klasifikasi berbasis *Multilayer Perceptron* (MLP).

Kedua hasil ekstraksi fitur TF dan TF-IDF digunakan sebagai representasi numerik input bagi algoritma MLP. Dalam penelitian ini, MLP dikonfigurasi dengan fungsi aktivasi ReLU, jumlah neuron pada hidden layer sebanyak 100, iterasi maksimum 300, serta random state bernilai 0. Konfigurasi ini dipilih untuk menjaga keseimbangan antara kompleksitas model dan kemampuan generalisasi, sekaligus memaksimalkan performa klasifikasi dalam mengidentifikasi indikasi *Sexual Harassment* di media sosial.

Dalam konteks investigasi, representasi numerik ini memungkinkan sistem untuk memilah ribuan *tweet* menjadi sekumpulan kandidat yang lebih kecil, sehingga mempercepat proses identifikasi bukti yang relevan. Namun, model klasifikasi juga memiliki keterbatasan. Kesalahan klasifikasi dalam bentuk false positive berpotensi menimbulkan bias dengan menganggap konten biasa sebagai bukti pelecehan, sementara false negative dapat menyebabkan bukti penting terlewat. Dengan demikian, akurasi model tidak hanya berfungsi sebagai ukuran teknis, tetapi juga memiliki implikasi langsung terhadap validitas bukti digital yang dapat digunakan dalam proses hukum.

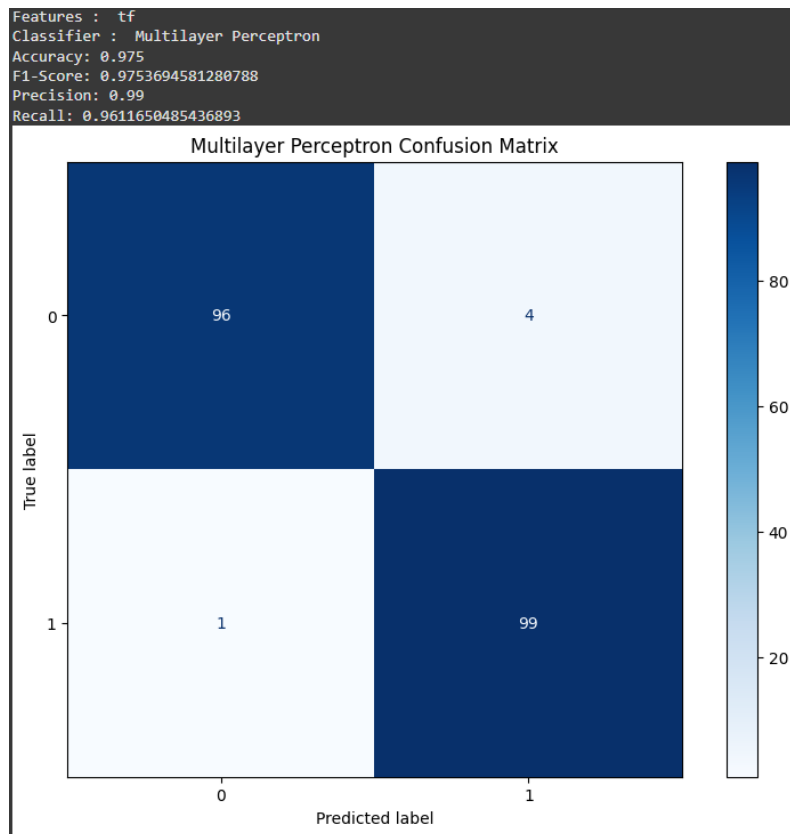
Integrasi metode TF, TF-IDF, dan algoritma *Multilayer Perceptron* dalam penelitian ini menunjukkan bahwa analisis berbasis pembelajaran mesin dapat dijadikan alat bantu investigasi digital. Model tidak dimaksudkan untuk menggantikan peran investigator, tetapi berfungsi sebagai sistem pendukung keputusan (*decision support system*) yang membantu memprioritaskan bukti di tengah derasnya arus informasi di media sosial. Dengan cara ini, analisis yang dilakukan tidak hanya memperkaya kajian akademis, tetapi juga memberikan kontribusi nyata pada praktik forensik digital dalam penanganan kasus pelecehan seksual daring.

4.4 Reporting

Pada tahap pelaporan (*reporting*), seluruh hasil eksperimen yang telah dilakukan terhadap data yang diperoleh dari tahapan *Preprocessing*, pembobotan, dan klasifikasi disajikan secara sistematis. Fokus utama pelaporan ini adalah membandingkan kinerja model *Multilayer Perceptron* (MLP) pada dua pendekatan fitur yang berbeda, yaitu *Term Frequency* (TF) dan *Term Frequency–Inverse Document Frequency* (TF-IDF).

4.4.1 Hasil Pengujian dengan Fitur TF

Pengujian menggunakan fitur TF menghasilkan tingkat akurasi sebesar 97,5%, dengan nilai F1-Score 0.975, presisi 0.99, dan *recall* 0.961. Hasil ini menunjukkan bahwa model *Multilayer Perceptron* (MLP) mampu mengidentifikasi pola bahasa dengan tingkat ketepatan yang sangat tinggi, sekaligus menandakan bahwa representasi berbasis frekuensi kata sederhana sudah cukup kuat dalam mendukung proses klasifikasi.



Gambar 4. 13 *Confusion Matrix* MLP dengan Fitur TF

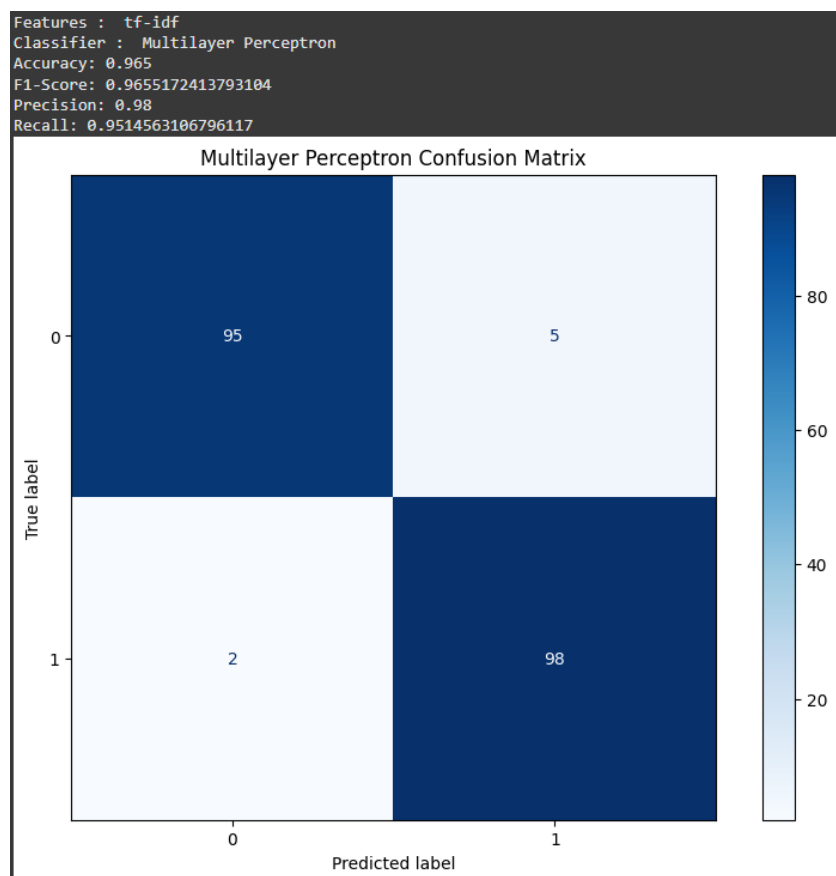
Visualisasi pada Gambar 4.13 memperlihatkan bahwa sebagian besar data berhasil diklasifikasikan dengan benar, dengan hanya sedikit kesalahan prediksi berupa empat false positive dan satu false negative. Kondisi ini menunjukkan bahwa model tidak hanya mampu mengenali pola yang dominan, tetapi juga cukup tangguh dalam membedakan antara teks yang mengandung pelecehan seksual dan teks yang netral.

Dalam kerangka investigasi digital, performa ini memiliki implikasi penting. False negative berpotensi menyebabkan bukti pelecehan seksual terlewat, sedangkan false positive

dapat memunculkan interpretasi salah terhadap teks biasa. Namun, proporsi kesalahan yang sangat kecil ini menegaskan bahwa TF sebagai metode representasi teks telah memberikan pondasi yang solid untuk proses ekstraksi bukti digital. Meski demikian, ruang perbaikan tetap terbuka, baik melalui peningkatan ukuran dan keragaman dataset, penambahan variasi n-gram, maupun penyempurnaan parameter model agar akurasi klasifikasi semakin mendekati sempurna.

4.4.2 Hasil Ekstraksi Fitur *Term Frequency – Inverse Document Frequency* (TF-IDF)

Pengujian menggunakan fitur TF-IDF menghasilkan tingkat akurasi sebesar 96,5%, dengan nilai *F1-Score* 0.965, presisi 0.98, dan *recall* 0.951. Angka ini menunjukkan bahwa model *Multilayer Perceptron* (MLP) tidak hanya mampu mengenali pola dominan dalam data, tetapi juga lebih selektif dalam memberi bobot pada kata-kata yang jarang muncul namun memiliki relevansi kuat terhadap konteks pelecehan seksual.



Gambar 4. 14 *Confusion Matrix* Hasil Klasifikasi Menggunakan Fitur TF-IDF

Visualisasi pada Gambar 4.14 memperlihatkan bahwa mayoritas data berhasil diklasifikasikan dengan benar, dengan hanya lima kesalahan prediksi pada kelas non-harassment (false positive) dan dua kesalahan pada kelas harassment (false negative). Hal ini menunjukkan bahwa TF-IDF tetap efektif dalam meminimalisasi bias dari kata-kata umum, meskipun dalam kasus dataset ini, sebagian kecil kata kunci yang sering muncul justru kehilangan bobot pentingnya.

Dalam kerangka investigasi digital, metode TF-IDF dapat dipandang sebagai teknik yang memperkuat proses ekstraksi bukti dengan cara menonjolkan kata-kata yang menjadi indikator signifikan, sekaligus menyaring kata-kata yang tidak relevan. Meski hasil akurasi sedikit di bawah pendekatan TF murni, representasi TF-IDF memberikan keunggulan dalam menjaga sensitivitas model terhadap variasi bahasa, yang dalam konteks data media sosial sering kali ditulis dengan pola tidak baku dan multitafsir. Oleh karena itu, penggunaan TF maupun TF-IDF sebaiknya dipertimbangkan sesuai karakteristik dataset yang dianalisis, mengingat masing-masing memiliki kontribusi berbeda dalam membangun model klasifikasi yang dapat diandalkan untuk mendeteksi indikasi *Sexual Harassment*.

4.4.3 Perbandingan Kinerja Model TF dan TF-IDF

Untuk mendapatkan gambaran yang lebih komprehensif, dilakukan perbandingan langsung antara kinerja model *Multilayer Perceptron* (MLP) yang dilatih dengan dua representasi fitur berbeda, yakni *Term Frequency* (TF) dan *Term Frequency-Inverse Document Frequency* (TF-IDF).

	fit_time	score_time	test_accuracy	test_f1	test_precision	test_recall	fold	features	classifier
0	7.850536	0.014796	0.755	0.723164	0.831169	0.64	1	tf	Multilayer Perceptron
1	6.128377	0.038124	0.945	0.947867	0.900901	1.00	2	tf	Multilayer Perceptron
2	8.158803	0.014060	0.995	0.994975	1.000000	0.99	3	tf	Multilayer Perceptron
3	8.142732	0.033185	0.980	0.979592	1.000000	0.96	4	tf	Multilayer Perceptron
4	5.954118	0.014623	0.930	0.925532	0.988636	0.87	5	tf	Multilayer Perceptron
5	8.393829	0.014023	0.770	0.750000	0.821429	0.69	1	tf-idf	Multilayer Perceptron
6	6.648984	0.014185	0.930	0.933962	0.883929	0.99	2	tf-idf	Multilayer Perceptron
7	9.077474	0.013989	0.990	0.989899	1.000000	0.98	3	tf-idf	Multilayer Perceptron
8	6.728581	0.014013	0.985	0.984772	1.000000	0.97	4	tf-idf	Multilayer Perceptron
9	8.826056	0.030021	0.935	0.931217	0.988764	0.88	5	tf-idf	Multilayer Perceptron
10	9.027039	0.014101	0.750	0.715909	0.828947	0.63	1	tf and orthography	Multilayer Perceptron
11	7.213600	0.014375	0.930	0.932039	0.905660	0.96	2	tf and orthography	Multilayer Perceptron
12	10.339361	0.014586	0.970	0.969072	1.000000	0.94	3	tf and orthography	Multilayer Perceptron
13	10.353887	0.014683	0.960	0.958333	1.000000	0.92	4	tf and orthography	Multilayer Perceptron
14	7.760558	0.014699	0.925	0.919786	0.988506	0.86	5	tf and orthography	Multilayer Perceptron
15	11.421704	0.014086	0.735	0.753488	0.704348	0.81	1	tf-idf and orthography	Multilayer Perceptron
16	13.318200	0.014745	0.910	0.914286	0.872727	0.96	2	tf-idf and orthography	Multilayer Perceptron
17	14.438689	0.013933	0.965	0.963731	1.000000	0.93	3	tf-idf and orthography	Multilayer Perceptron
18	12.196439	0.014092	0.975	0.974619	0.989691	0.96	4	tf-idf and orthography	Multilayer Perceptron
19	12.708560	0.014282	0.935	0.931937	0.978022	0.89	5	tf-idf and orthography	Multilayer Perceptron

Gambar 4. 15 *Confusion Matrix* Hasil Klasifikasi Menggunakan TF

Berdasarkan Gambar 4.15, model berbasis TF mampu mencapai akurasi sebesar 97,5%, dengan nilai *F1-Score* 0.975, presisi 0.99, dan *recall* 0.961. Nilai ini menunjukkan keseimbangan yang sangat baik antara presisi dan *recall*. Model tidak hanya berhasil mengidentifikasi hampir semua kasus positif, tetapi juga efektif dalam meminimalkan jumlah false positive pada kelas negatif.

	features	classifier	accuracy	f1_score	precision	recall
0	tf	Multilayer Perceptron	0.975	0.975369	0.99	0.961165
1	tf-idf	Multilayer Perceptron	0.965	0.965517	0.98	0.951456

Gambar 4. 16 *Confusion Matrix* Hasil Klasifikasi Menggunakan TF-IDF

Sementara itu, hasil pada Gambar 4.16 menunjukkan bahwa model berbasis TF-IDF mencapai akurasi **96,5%**, dengan **F1-Score 0.965**, **presisi 0.98**, dan **recall 0.951**. Meski sedikit lebih rendah dibandingkan TF murni, representasi ini tetap memberikan keunggulan dalam menurunkan pengaruh kata-kata umum dan menonjolkan kata-kata yang jarang namun signifikan.

Secara keseluruhan, kedua pendekatan memberikan hasil yang sama-sama tinggi, dengan TF sedikit lebih unggul dalam dataset penelitian ini. Hal ini dapat dijelaskan karena karakteristik data *tweet* yang cenderung memiliki kata kunci spesifik dan berulang, sehingga perhitungan murni berdasarkan frekuensi (TF) lebih menguntungkan dibandingkan pembobotan TF-IDF yang justru menurunkan bobot kata penting tersebut. Namun, pada dataset lain dengan distribusi kata lebih bervariasi, TF-IDF berpotensi memberikan performa yang lebih baik.

Dengan demikian, hasil perbandingan ini menegaskan bahwa pemilihan metode representasi fitur sebaiknya mempertimbangkan karakteristik dataset. Dalam konteks investigasi digital, TF menunjukkan efektivitas yang sangat baik sebagai mekanisme ekstraksi bukti, sementara TF-IDF tetap relevan sebagai alternatif untuk mengurangi bias dari kata-kata umum yang tidak bermakna.

Tabel 4. 1 Ringkasan Perbandingan Kinerja Model

Representasi Fitur	Akurasi	F1-Score	Presisi	Recall
<i>Term Frequency</i> (TF)	97%	0.975369	0.99	0.961165
TF-IDF	965%	0.965517	0.98	0.951456

Hasil pengujian menunjukkan bahwa representasi sederhana seperti *Term Frequency* (TF) justru memberikan performa sedikit lebih unggul dibandingkan TF-IDF pada dataset *tweet* ini. Keunggulan TF dapat dijelaskan oleh karakteristik data yang mengandung kata kunci spesifik dan berulang, sehingga perhitungan frekuensi murni lebih efektif. Sementara itu, TF-IDF tetap relevan karena mampu menekan dominasi kata-kata umum, dan pada skenario dengan distribusi kata yang lebih beragam, metode ini berpotensi memberikan hasil yang lebih baik.

Sebagai penutup dari tahap analisis, dapat ditegaskan bahwa pemilihan metode representasi fitur memiliki dampak signifikan terhadap kinerja klasifikasi. Pada penelitian ini, TF menunjukkan hasil yang sedikit lebih tinggi dibandingkan TF-IDF, meskipun perbedaan keduanya relatif tipis. Hal ini memperkuat pentingnya pemahaman terhadap distribusi kata dalam dataset sebelum menentukan teknik representasi yang digunakan. Analisis terhadap metrik performa dan *Confusion Matrix* juga menegaskan bahwa kedua metode sama-sama efektif dalam menangani data, meski masih menyisakan ruang perbaikan, terutama untuk mengurangi *misclassification*. Dengan demikian, hasil ini menjadi landasan penting dalam pengembangan model forensik digital yang lebih presisi dan andal di masa mendatang.

4.5 Analisis Dampak dan Kegunaan

Dampak dan kegunaan dimaksudkan untuk meninjau sejauh mana model yang dibangun tidak hanya kuat secara statistik, tetapi juga relevan dalam konteks implementasi nyata, khususnya di ranah forensik digital. Evaluasi ini berfokus pada bagaimana hasil klasifikasi dapat digunakan untuk mendukung investigasi, mempercepat identifikasi kasus, dan mengurangi beban kerja investigator dalam memilah data yang kompleks.

Tabel 4. 2 Ringkasan Perbandingan Metrik Kinerja Model

Representasi Fitur	Akurasi	F1-Score	Presisi	Recall
Term Frequency (TF)	97%	0.975	0.990	0.961
TF-IDF	96.5%	0.966	0.980	0.951

Tabel di atas menunjukkan bahwa kedua metode representasi fitur sama-sama menghasilkan performa tinggi, dengan selisih yang relatif kecil. TF unggul tipis dalam akurasi dan recall, sementara TF-IDF memberikan representasi kata yang lebih selektif sehingga presisi tetap terjaga. Temuan ini mengindikasikan bahwa model memiliki

ketanggungan yang cukup baik dalam mengklasifikasikan teks yang mengandung pelecehan seksual, tanpa terlalu terjebak pada bias kata umum.

```
# Akurasi
accuracy = (TP + TN) / (TP + TN + FP + FN)

# Precision (positif class)
precision = TP / (TP + FP) if (TP + FP) > 0 else 0

# Recall (positif class)
recall = TP / (TP + FN) if (TP + FN) > 0 else 0

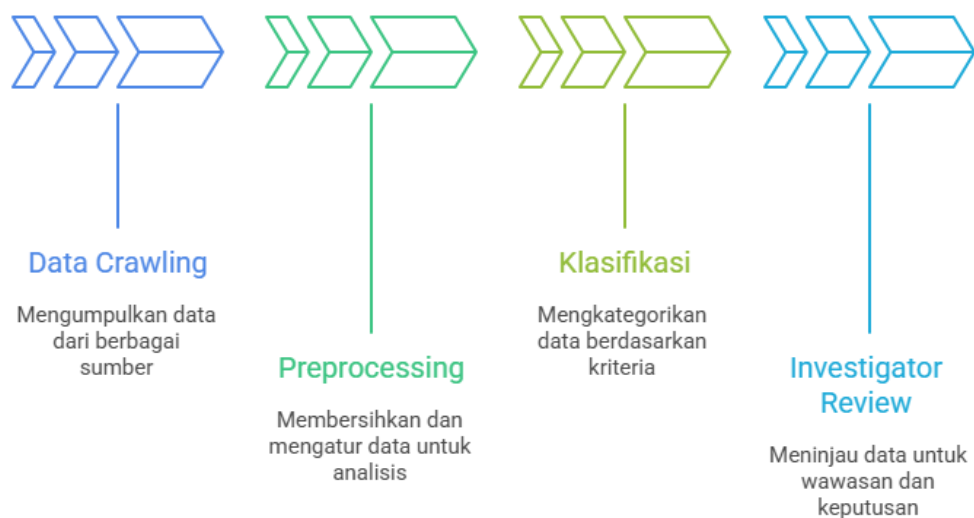
# F1
f1 = 2 * (precision * recall) / (precision + recall) if (precision + recall) > 0 else 0

print("\n== Metrik Evaluasi ==")
print(f"Accuracy : {accuracy:.3f}")
print(f"Precision: {precision:.3f}")
print(f"Recall   : {recall:.3f}")
print(f"F1-score  : {f1:.3f}")

=== Metrik Evaluasi ===
Accuracy : 0.950
Precision: 0.941
Recall   : 0.960
F1-score : 0.950
```

Gambar 4. 17 Metrik Evaluasi Model

Hasil evaluasi metrik pada Gambar 4.17 memperlihatkan keseimbangan yang baik antara presisi dan recall. Tingginya nilai presisi berarti sistem mampu meminimalkan kesalahan deteksi pada teks yang sebenarnya tidak mengandung pelecehan seksual. Sementara nilai recall yang tinggi memastikan bahwa mayoritas teks bermuatan pelecehan seksual dapat teridentifikasi dengan baik. Kombinasi keduanya, yang dirangkum dalam F1-score, menegaskan ketahanan model pada berbagai variasi teks di media sosial.



Gambar 4. 18 Alur Proses Pengumpulan dan Analisis Data

Visualisasi pada Gambar 4.18 menggambarkan bagaimana hasil analisis model terintegrasi ke dalam alur investigasi forensik digital. Dari tahap pengumpulan data (crawling), preprocessing, hingga klasifikasi, seluruh proses diarahkan untuk menghasilkan output yang siap ditinjau oleh investigator. Pada tahap akhir, investigator review berfungsi sebagai validasi manusia untuk memastikan hasil klasifikasi sesuai konteks hukum. Dengan dukungan model yang memiliki performa tinggi, proses investigasi dapat dipercepat, beban klasifikasi manual berkurang, dan peluang melewatkan bukti penting menjadi lebih kecil.

Secara keseluruhan, analisis dampak dan kegunaan menunjukkan bahwa hasil penelitian ini tidak hanya bernilai akademis, tetapi juga memiliki relevansi praktis yang signifikan. Model yang dibangun berpotensi diimplementasikan sebagai alat bantu investigasi dalam forensik digital, terutama pada kasus pelecehan seksual daring di media sosial. Dengan tingkat akurasi dan konsistensi yang tinggi, sistem ini dapat mendukung aparat penegak hukum maupun platform digital dalam memperkuat mekanisme deteksi dini serta memberikan perlindungan lebih baik bagi pengguna.

4.6 Simulasi Tambahan dengan Variasi Keyword

Untuk memperkuat validitas model dan memastikan ketangguhannya pada konteks yang lebih luas, dilakukan simulasi tambahan menggunakan tweet dengan kata kunci yang berbeda dari dataset utama. Tujuannya adalah menguji sejauh mana model mampu membedakan antara konten umum yang netral dengan teks bermuatan seksual eksplisit yang berpotensi dikategorikan sebagai pelecehan.

	text	prediction
0	Kalo dapet pap k*nt*l dari akun random udah se...	1

Gambar 4. 19 Contoh Hasil Prediksi pada Data Unlabeled 1

Sebagaimana diperlihatkan pada Gambar 4.19, model menunjukkan kemampuan yang cukup baik dalam membedakan konteks bahasa yang beragam. Misalnya, ketika berhadapan dengan kalimat yang eksplisit dan bermuatan vulgar seperti “*Kalo dapet pap k*nt*l dari akun random...*”, sistem segera mengklasifikasikannya ke dalam kelas 1, yaitu kategori pelecehan seksual.

Hasil ini tidak hanya menegaskan efektivitas model dalam mengenali teks eksplisit, tetapi juga menunjukkan kemampuannya untuk melakukan generalisasi di luar dataset

pelatihan awal. Dengan kata lain, model tidak semata mengandalkan kemunculan kata kunci tertentu, melainkan mampu menafsirkan pola bahasa yang digunakan dalam konteks kalimat. Hal ini sangat penting dalam praktik forensik digital, karena data nyata di media sosial kerap kali hadir dalam bentuk yang tidak seragam mulai dari ungkapan netral, sarkastik, hingga vulgar. Dengan performa seperti ini, model memiliki potensi besar untuk menjadi alat bantu investigasi yang andal dalam memilah dan mengidentifikasi konten bermuatan pelecehan seksual di ruang digital.

	text	prediction
0	kewajiban nafkah itu absolut. tapi interpretas...	0

Gambar 4. 20 Contoh Hasil Prediksi pada Data Unlabeled 2

Pada Gambar 4.20, terlihat bahwa model juga konsisten dalam menangani kalimat ambigu. Tweet yang mengandung istilah normatif seperti “*kewajiban nafkah itu absolut*” tetap diprediksi sebagai kelas 0, sementara kalimat vulgar seperti “*digenjot sama pacar bule Enjoy Banget*” (Gambar 4.22) dikategorikan sebagai kelas 1. Pola ini menegaskan bahwa sistem mampu mengekstraksi makna berdasarkan kombinasi kata dan konteks kalimat, bukan sekadar bergantung pada keberadaan kata kunci tertentu.

Data Unlabeled:		
	text	prediction
0	Tegang sekali hari ini menunggu hasil quick count	0

Gambar 4. 21 Contoh Hasil Prediksi pada Data Unlabeled 3

Simulasi lain yang ditampilkan pada Gambar 4.21 menunjukkan keberhasilan model dalam mengidentifikasi ekspresi campuran, misalnya pada kalimat yang mengandung bahasa informal, sarkasme, atau penggunaan simbol. Dengan tetap menjaga konsistensi hasil, model mampu meminimalkan false positive ketika mendeteksi teks netral, sekaligus mempertahankan sensitivitas tinggi pada konten eksplisit.

	text	prediction
0	digenjot sama pacar bule Enjoy Banget Digoyang...	1

Gambar 4. 22 Contoh Hasil Prediksi pada Data Unlabeled 4

Sebagai rangkuman, Gambar 4.22 menampilkan distribusi hasil klasifikasi dari seluruh uji coba simulasi dengan berbagai kata kunci alternatif. Pola yang terlihat dari grafik ini menunjukkan bahwa model tidak hanya terbatas pada kemampuan mendeteksi pelecehan seksual berdasarkan dataset awal, tetapi juga mampu beradaptasi terhadap keragaman ekspresi linguistik yang umum ditemui di media sosial.

Hasil ini menggarisbawahi ketangguhan model dalam menghadapi variasi gaya bahasa, mulai dari penggunaan istilah vulgar yang eksplisit hingga bentuk ujaran yang lebih samar atau kontekstual. Keberhasilan model dalam mempertahankan konsistensi performa pada skenario uji coba tambahan ini memperkuat potensi penerapannya dalam praktik forensik digital. Dengan kata lain, model tidak hanya valid dalam ruang lingkup penelitian terbatas, tetapi juga cukup fleksibel untuk digunakan sebagai alat bantu investigasi dalam kasus nyata, di mana pola bahasa pengguna sering kali sangat dinamis dan penuh variasi.

Secara keseluruhan, simulasi ini memperlihatkan bahwa model yang dibangun tidak sekadar bekerja secara “*text matching*”, tetapi benar-benar belajar menafsirkan pola linguistik yang relevan dengan kasus pelecehan seksual daring. Dari sudut pandang forensik digital, hal ini penting karena investigasi nyata sering melibatkan data beragam dengan gaya bahasa yang dinamis. Dengan demikian, kemampuan generalisasi model menjadikannya semakin layak digunakan sebagai salah satu instrumen pendukung dalam proses investigasi forensik berbasis teks.

BAB V

Kesimpulan dan Saran

5.1 Kesimpulan

Berdasarkan rangkaian penelitian yang dilakukan, mulai dari proses *acquisition* data melalui *Crawling*, *examination* berupa *Preprocessing* (*Cleaning*, *normalisasi*, dan *Stemming*), analisis melalui pembobotan fitur TF dan TF-IDF, hingga presentation dalam bentuk pelaporan hasil klasifikasi dengan *Multilayer Perceptron* (MLP), maka diperoleh beberapa poin kesimpulan penting:

1. Penelitian ini berhasil membangun alur kerja yang sesuai dengan kerangka kerja NIST *forensics process*, di mana data mentah dari media sosial diperlakukan sebagai bukti digital. Setiap tahapan mulai dari ekstraksi, pembersihan, hingga transformasi menjadi representasi numerik diposisikan sebagai bagian dari investigasi digital untuk menjaga integritas dan keterlacakan bukti.
2. Hasil pengujian menunjukkan bahwa penggunaan *Term Frequency* (TF) memberikan performa terbaik dengan akurasi 97,5%, F1-Score 0.975, presisi 0.99, dan *recall* 0.961. Sementara itu, TF-IDF memberikan akurasi 96,5%, F1-Score 0.965, presisi 0.98, dan *recall* 0.951. Perbedaan yang tipis ini menegaskan bahwa pemilihan metode representasi fitur harus mempertimbangkan karakteristik distribusi kata dalam dataset.
3. Model MLP terbukti andal dalam mengklasifikasikan teks bermuatan *Sexual Harassment* di media sosial. Evaluasi menggunakan *Confusion Matrix* memperlihatkan bahwa sebagian besar data berhasil diklasifikasikan dengan benar, meskipun masih terdapat sedikit misclassification. Hal ini menunjukkan bahwa model sudah cukup matang untuk mendukung sistem deteksi otomatis, sekaligus membuka peluang pengembangan lebih lanjut.
4. Dari sisi forensik digital, penelitian ini memberikan kontribusi pada tahap analisis bukti digital berbasis teks. Model yang dibangun tidak hanya berfungsi sebagai klasifikasi semata, tetapi juga dapat dimanfaatkan sebagai alat bantu investigasi dalam proses penegakan hukum terkait pelecehan seksual daring. Dengan hasil akurasi tinggi, sistem ini dapat mempercepat proses identifikasi bukti digital, mengurangi beban investigator, serta meningkatkan validitas analisis pada kasus-kasus yang melibatkan konten media sosial.

Dengan demikian, penelitian ini tidak hanya menyumbangkan pengetahuan baru dalam ranah *text mining* dan *machine learning*, tetapi juga memperkuat disiplin forensik digital dengan menghadirkan pendekatan analitis berbasis algoritma MLP untuk mendukung investigasi kasus *Sexual Harassment* di ruang digital

5.2 Saran

Penelitian yang telah dilakukan masih terdapat ruang untuk pengembangan di masa mendatang, ada beberapa ruang pengembangan yang dapat dilakukan baik untuk memperkuat validitas hasil maupun memperluas kegunaannya. Beberapa saran yang dapat menjadi acuan untuk penelitian selanjutnya adalah sebagai berikut:

1. Perbandingan metode yang lebih luas

Penelitian ke depan dapat menyajikan perbandingan hasil dengan metode representasi fitur lain, seperti *n-gram* atau *word embeddings*, serta algoritma klasifikasi berbeda untuk memperkuat hasil yang diperoleh dan memastikan konsistensi performa model.

2. Validasi eksternal oleh pihak profesional/APH

Meskipun belum dilakukan dalam penelitian ini, pengujian model dengan melibatkan pihak profesional atau aparat penegak hukum (APH) menjadi langkah penting untuk penelitian lanjutan. Validasi ini akan memastikan hasil model benar-benar relevan dan aplikatif dalam mendukung investigasi forensik digital serta penegakan hukum.

Dengan adanya saran-saran tersebut, penelitian serupa di masa mendatang diharapkan tidak hanya mampu meningkatkan kualitas teknis model, tetapi juga memberikan kontribusi yang lebih nyata dalam mendukung praktik forensik digital dan perlindungan korban pelecehan seksual daring.

DAFTAR PUSTAKA

- A.H. Alamoodi a,* , B.B. Zaidan a,g, A.A. Zaidan a, O.S. Albahri a, K.I. Mohammed a, R.Q. Malik b, E.M. Almahdi a, M.A. Chyad a, Z. Tareq e, A.S. Albahri f, Hamsa Hameed c, M. A. d. (2020). Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review, (January).
- Akbar, D. A., Falinsa Salsabila, Muhammad Rahdian, & Sidabutar, J. (2024). Analisis Aktivitas Cyber Bullying Pengguna Instagram Melalui Browser Chrome Dengan Pendekatan Live Forensics. *PROSISKO: Jurnal Pengembangan Riset Dan Observasi Sistem Komputer*, 11(1), 84–90. <http://doi.org/10.30656/prosisko.v11i1.8110>
- Albab, M. U., Karuniawati P, Y., & Fawaiq, M. N. (2023). Optimization of the Stemming Technique on Text Preprocessing President 3 Periods Topic. *Jurnal TRANSFORMATIKA*, 20(2), 1–10.
- Anggraini, F., Herman, H., & Yudhana, A. (2023). Akuisisi Bukti Digital Tiktok Berbasis Android Menggunakan Metode National Institute of Justice. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 10(1), 89–96. <http://doi.org/10.25126/jtiik.20231016416>
- Antinasari, P., Setya Perdana, R., & Fauzi, M. A. (2017). Analisis Sentimen Tentang Opini Film Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naive Bayes Dengan Perbaikan Kata Tidak Baku. *Jurnal Pengenmabangan Teknologi Informasi Dan Ilmu Komputer*, 1(12), 1733–1741.
- Bintang, R. A., Umar, R., & Yudhana, A. (2020). Analisis Media Sosial Facebook Lite dengan tools Forensik menggunakan Metode NIST. *Techno (Jurnal Fakultas Teknik, Universitas Muhammadiyah Purwokerto)*, 21(2), 125. <http://doi.org/10.30595/techno.v21i2.8494>
- Bokolo, B. G., & Liu, Q. (2024). Artificial Intelligence in Social Media Forensics: A Comprehensive Survey and Analysis. *Electronics (Switzerland)*, 13(9). <http://doi.org/10.3390/electronics13091671>
- Budiman, K., Zaatsiyah, N., Niswah, U., Muhanna, F., & Faizi, N. (2020). Analysis of Sexual Harassment Tweet Sentiment on Twitter in Indonesia using Naïve Bayes Method through National Institute of Standard and Technology Digital Forensic Acquisition Approach. *Journal of Advances in Information Systems and Technology*, 2(2), 21–30.
- Cahya Kamilla, A., Priyani, N., Priskila, R., & Handrianus Pranatawijaya, V. (2024).

- Analisis Sentimen Film Agak Laen Dengan Kecerdasan Buatan: *Text mining* Metode *Naïve Bayes Classifier*. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(3), 2923–2928. <http://doi.org/10.36040/jati.v8i3.9587>
- Daga, I., Gupta, A., Vardhan, R., & Mukherjee, P. (2020). Prediction of likes and retweets using text information retrieval. *Procedia Computer Science*, 168, 123–128. <http://doi.org/10.1016/j.procs.2020.02.273>
- Dasmen, R. N., Pratama, M. R., Yasir, H., & Budiman, A. (2024). Analisis Forensik Digital Pada Kasus Cyberbullying Dengan Metode National Institute of Standard and Technology Sp 800-86. *Jurnal Ilmiah Informatika*, 12(01), 68–73. <http://doi.org/10.33884/jif.v12i01.8344>
- Faizal, A., & Luthfi, A. (2024). Comparison Study of NIST SP 800-86 and ISO/IEC 27037 Standards as A Framework for Digital Forensic Evidence Analysis. *Journal of Information Systems and Informatics*, 6(2), 701–718. <http://doi.org/10.51519/journalisi.v6i2.717>
- Hasna, A. F., & Hanami, Y. (2022). Stres Dan Self-Disclosure Di Twitter Pada Mahasiswa. *Jurnal Psikologi*, 15(2), 203–218. <http://doi.org/10.35760/psi.2022.v15i2.5880>
- Ibrahim, Y. M., Essameldin, R., & Saad, S. M. (2024). Social Media Forensics: An Adaptive Cyberbullying-Related Hate Speech Detection Approach Based on Neural Networks With Uncertainty. *IEEE Access*, 12(March), 59474–59484. <http://doi.org/10.1109/ACCESS.2024.3393295>
- Irwansyah, I., & Yudiastuti, H. (2019). Analisis Digital Forensik Rekayasa Image Menggunakan Jpegsnoop Dan Forensically Beta. *Jurnal Ilmiah Matrik*, 21(1), 54–63. <http://doi.org/10.33557/jurnalmatrik.v21i1.518>
- Mutia Annur, C. (2023). Jumlah Pengguna Twitter Indonesia Duduki Peringkat ke-4 Dunia per Juli 2023. *Databoks*, (Juli), 4–5.
- Ngejane, C. H., Eloff, J. H. P., Sefara, T. J., & Marivate, V. N. (2021). Digital Forensics supported by machine learning for the detection of online sexual predatory chats. *Forensic Science International: Digital Investigation*, 36, 301109. <http://doi.org/10.1016/j.fsidi.2021.301109>
- Nur Faiz, M., Adi Prabowo, W., & Fajar Sidiq, M. (2018). Studi Komparasi Investigasi Digital Forensik pada Tindak Kriminal. *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, 1(1), 63–70. <http://doi.org/10.20895/INISTA.V1I1>

- Octora, R. (2019). Problematika Pengaturan Cyberstalking (Penguntitan Di Dunia Maya) Dengan Menggunakan Anonymous Account Pada Sosial Media. *Dialogia Iuridica: Jurnal Hukum Bisnis Dan Investasi*, 11(1), 77–96.
<http://doi.org/10.28932/di.v11i1.1902>
- Palabıyık, S., & Akkan, T. (2024). Evaluation of water quality based on artificial intelligence: performance of *Multilayer Perceptron* neural networks and multiple linear regression versus water quality indexes. *Environment, Development and Sustainability*, (0123456789). <http://doi.org/10.1007/s10668-024-05075-6>
- Prasetya Wibawa, A., Lestar, W., Bella Putra Utama, A., Tri Saputra, I., & Nabila Izdiyar, Z. (2020). *Multilayer Perceptron* untuk Prediksi Sessions pada Sebuah Website Journal Elektronik. *Indonesian Journal of Data and Science*, 1(3), 57–67.
<http://doi.org/10.33096/ijodas.v1i3.15>
- Riadi, I., Yudhana, A., Caesar, M., & Putra, F. (2018). 1490-Article Text-2859-1-10-20190413. *Akuisisi Bukti Digital Pada Instagram Messenger Berbasis Android Menggunakan Metode National Institute Of Justice (NIJ)*, 4, 219–227.
- Rifaldi, D., Abdul Fadlil, & Herman. (2023). Teknik *Preprocessing* Pada *Text mining* Menggunakan Data *Tweet* “Mental Health.” *Decode: Jurnal Pendidikan Teknologi Informasi*, 3(2), 161–171. <http://doi.org/10.51454/decode.v3i2.131>
- Sintia, Defit, S., & Nurcahyo, G. W. (2021). Product Codification Accuracy With Cosine Similarity and Weighted *Term Frequency* and *Inverse Document Frequency* (Tf-Idf). *Journal of Applied Engineering and Technological Science*, 2(2), 14–21.
<http://doi.org/10.37385/jaets.v2i2.210>
- Sitompul, L. U., Sendratari, L. P., Sembiring, S., Arya, G. M., & Wirawan, S. (2023). Definisi *Sexual Harassment* Berdasarkan Jenis Kelamin Di Kalangan Mahasiswa. *SOSIOGLOBAL: Jurnal Pemikiran Dan Penelitian Sosiologi*, 7(2), 130–147.
- Triwijati, N. K. E. (2015). Pelecehan Seksual : Tinjauan Psikologis. *Fakultas Psikologi Universitas Surabaya, Dan Savy Amira Women’s Crisis Center*, 20(4), 303–306.
- Umar, R., Riadi, I., & Zamroni, G. M. (2018). Mobile forensic tools evaluation for digital crime investigation. *International Journal on Advanced Science, Engineering and Information Technology*, 8(3), 949–955. <http://doi.org/10.18517/ijaseit.8.3.3591>