KLASIFIKASI *CYBERBULLYING* MENGGUNAKAN METODE *DESICION TREE* DENGAN EKSTRAKSI FITUR *TERM FREQUENCY-INVERCE DOCUMENT* (TF-IDF)

SKRIPSI

Oleh:
SARI INDAH LESTARI
NIM. 19650045

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026

KLASIFIKASI *CYBERBULLYING* MENGGUNAKAN METODE *DESICION TREE* DENGAN EKSTRAKSI FITUR *TERM FREQUENCY-INVERCE DOCUMENT* (TF-IDF)

SKRIPSI

Diajukan Kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh :
SARI INDAH LESTARI
NIM. 19650045

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2026

HALAMAN PERSETUJUAN

KLASIFIKASI *CYBERBULLYING* MENGGUNAKAN METODE *DECISION TREE* DENGAN EKSTRAKSI FITUR *TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY* (TF-IDF)

SKRIPSI

Oleh :
SARI INDAH LESTARI
NIM. 19650045

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 24 Juni 2026

Pembimbing I,	Pembimbing II,
Okta Qomaruddin Aziz, M.Kom	Johan Ericka Wahyu Prakasa, M.Kom
NIP. 19911019 201903 1 013	NIP. 19831213 201903 1 004

Mengetahui,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Supriyono, M. Kom
NIP. 19841010 201903 1 012

HALAMAN PENGESAHAN

KLASIFIKASI *CYBERBULLYING* MENGGUNAKAN METODE *DECISION TREE* DENGAN EKSTRAKSI FITUR *TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY* (TF-IDF)

SKRIPSI

Oleh :
SARI INDAH LESTARI
NIM. 19650045

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 25 Juni 2026

Susunan Dewan Penguji

Ketua Penguji	: Khadijah Fahmi Hayati Holle, M.Kom NIP. 19900626 202203 2 002	( )
Anggota Penguji I	: Fatchurrohman, M.Kom NIP. 19700731 200501 1 002	( )
Anggota Penguji II	: Okta Qomaruddin Aziz, M.Kom NIP. 19911019 201903 1 013	( )
Anggota Penguji III	: Johan Ericka Wahyu Prakasa, M.Kom NIP. 19831213 201903 1 004	( )

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim
Malang

Supriyono, M. Kom
NIP. 19841010 201903 1 012

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama	: Sari Indah Lestari
NIM	: 19650045
Fakultas / Program Studi	: Sains dan Teknologi / Teknik Informatika
Judul Skripsi	: PKlasifikasi *Cyberbullying* Menggunakan Metode *Decision Tree* dengan Ekstraksi Fitur *Term Frequency-Inverce Document* (TD-IDF)

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 22 Juni 2026

Yang membuat pernyataan,

Sari Indah Lestari

NIM. 19650045

MOTTO

If you wanna do, do it.
If you want to, get it

HALAMAN PERSEMBAHAN

Alhamdulillahi rabbil 'alamin segala puji bagi Allah SWT yang telah melimpahkan rahmat, hidayah, serta kemudahan sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Shalawat dan salam senantiasa tercurah kepada baginda Rasulullah SAW yang telah menuntun umat manusia keluar dari zaman jahiliyah menuju cahaya Islam.

Penulis mempersembahkan karya ini kepada kedua orang tua, diri sendiri, kakak tersayang, para dosen yang telah membimbing, teman-teman, serta orang-orang yang telah membantu, mendoakan, serta menyemangati penulis hingga selesainya skripsi ini.

KATA PENGANTAR

Segala puji bagi Allah Subhanahu wa Ta'ala, Tuhan semesta alam, atas karunia rahmat, nikmat, dan kesehatan yang telah diberikan sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Shalawat dan salam senantiasa kita panjatkan kepada Nabi Muhammad Shallallahu 'Alaihi Wasallam yang telah membawa agama penuh kebaikan hingga senantiasa bisa kita nikmati ini yaitu agama Islam.

Adapun penyusunan skripsi ini tidak lepas dari bantuan, dukungan, dan doa dari berbagai pihak, baik secara langsung maupun tidak langsung. Sebab itu, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. Hj. Ilfi Nur Diana, M.Si., CAHRM., CRMP selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang beserta seluruh jajaran pimpinan universitas.
2. Dr. Agus Mulyono, M.Si selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang beserta jajarannya.
3. Supriyono, M.Kom selaku Ketua Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Okta Qomaruddin Aziz, M.Kom selaku Dosen Pembimbing I dan Johan Ericka Wahyu Prakasa, M.Kom selaku Dosen Pembimbing II yang telah membimbing penulis dengan penuh kesabaran dan memberikan arahan selama proses penyusunan skripsi ini.

5. Fatchurrochman, M.Kom Khadijah Fahmi Hayati Holle, M.Kom selaku Ketua Penguji dan Khadijah Fahmi Hayati Holle, M.Kom selaku Penguji I yang telah memberikan masukan, kritik, dan saran yang sangat membangun demi penyempurnaan skripsi ini.
6. Seluruh Dosen dan Staf di Program Studi Teknik Informatika, dengan ikhlas memberikan ilmu, bantuan, serta dorongan semangat selama perkuliahan.
7. Kedua orang tua yang telah memberikan dukungan dan kepercayaan sampai terwujudnya skripsi ini.
8. Saudara dan teman-teman seangkatan 2019 yang telah memberikan motivasi dalam proses penulisan ini.
9. Semua pihak yang telah membantu dan memberikan dukungan dalam bentuk apapun, yang tidak dapat disebutkan satu per satu.

Dengan segala kerendahan hati penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, penulis berharap karya ini dapat memberikan manfaat, baik bagi pembaca maupun bagi penulis sendiri sebagai proses pembelajaran dan pengembangan diri.

Malang, 29 Juni 2026

Penulis

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

ABSTRAK

Lestari, Sari Indah. 2026. **Klasifikasi *Cyberbullying* Menggunakan Metode *Decision Tree* Dengan Ekstraksi Fitur *Term Frequency-Inverse Document Frequency* (TF-IDF)**. Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Okta Qomaruddin Aziz, M.Kom (II) Johan Ericka Wahyu Prakasa, M.Kom

Kata Kunci: Klasifikasi teks *cyberbullying, Decision Tree, Cyberbullying*

Seiring meningkatnya penggunaan jejaring media sosial oleh berbagai macam kalangan, *cyberbullying* bisa menjadi hal yang tidak bisa dihindari akan tetapi masih bisa disaring, meski demikian setiap individu juga harus selalu waspada terhadap *cyberbullying* dan diimbau untuk selalu melindungi diri ketika melakukan kegiatan di media sosial, sebab apabila dibiarkan saja tanpa penanganan *cyberbullying* dapat mengganggu kegiatan positif di media sosial. Dalam penelitan ini, penelitia akan memanfaatkan machine learning dalam mengklasifikasikan teks *cyberbullying* yang beredar pada teks tweet di twitter. Metode yang digunakan adalah *Decision Tree* untuk mengklasifikasikan data teks tweet sebanyak 47.692 data yang terbagi menjadi enam jenis kelas yakni, religon, gender, age, ethnicity, other cyberbullying, dan not cyberbullying. Setelah melalui tahapan pra proses data yakni pembersihan dan penyaringan teks, data nantinya akan dihitung bobotnya dengan TF-IDF, selanjutnya data akan diuji dengan tiga jenis skenario pembagian data, lalu diimplementasikan pada metode Decision Tree dengan kedalaman maksimum terbaik berdasarkan skenario data. Hasil yang diperoleh menujukkan skenario data 90:10 memiliki rata-rata akurasi dengan tinggi 83,63% dimana model mampu mengklasifikasikan data tesk *cyberbullying* berdasarkan kelas gender, age, ethnicity, religion, sedangkan model kurang dalam mengklasifikasi jenis kelas other cyberbullying dan not cyberbullying.

ABSTRACT

Lestari, Sari Indah. 2026. **Cyberbullying Classification Using the Decision Tree Method with Term Frequency-Inverse Document Frequency (TF-IDF) Feature Extraction**. *Thesis. Department of Informatics Engineering, Faculty of Science* and Technology, State Islamic University of Maulana Malik Ibrahim Malang. Advisors: (I) Okta Qomaruddin Aziz, M.Kom (II) Johan Ericka Wahyu Prakasa, M.Kom.

Keywords: Cyberbullying text classification, Decision Tree, Cyberbullying.

Along with the increasing use of social media networks by various groups, cyberbullying can become unavoidable but can still be filtered. Nevertheless, every individual must remain vigilant against cyberbullying and is urged to always protect themselves when engaging in social media activities, as if left unhandled, cyberbullying can disrupt positive activities on social media. In this study, the researcher utilizes machine learning to classify cyberbullying texts found in tweets on Twitter. The method used is the Decision Tree to classify 47,692 tweet text data, which are divided into six classes: religion, gender, age, ethnicity, other cyberbullying, and not cyberbullying. After going through the data preprocessing stages, namely text cleaning and filtering, the data weights are calculated using TF-IDF. Subsequently, the data is tested using three types of data splitting scenarios and then implemented into the Decision Tree method with the best maximum depth based on the data scenarios. The results obtained show that the 90:10 data scenario yields the highest mean accuracy of 83.63%, where the model is capable of classifying cyberbullying text data based on gender, age, ethnicity, and religion classes, while the model is less effective in classifying the other cyberbullying and not cyberbullying classes.

مستخلص البحث

ليستاري، ساري إنداه. ٢٠٢٦ .تصنيف التنمر السيبران باستخدام طريقة شجرة القرار بحث .*(TF-IDF)* مع استخراج الميزات بواسطة التردد اللفظي-التردد العكسي للوثيقة *(Decision Tree)* .جامعي .قسم هندسة المعلوماتية، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية بمالانق المشرفون) :١ (أوكتا قمر الدين عزيز، الماجستير ف علوم الحاسوب)٢ (جوهان إيريكا واهيو براكاسا، الماجستير ف .علوم الحاسوب

.الكلمات المفتاحية :تصنيف نصوص التنمر السيبراني، شجرة القرار، التنمر السيبراني

مع زيادة استخدام شبكات التواصل الاجتماعي من قبل مختلف الفئات، يمكن أن يصبح التنمر السيبران أمرًا لا مفر منه ولكنه لا يزال قابلًا للتصفية .ومع ذلك، يجب على كل فرد أن يظل يقظًا دائمًا ضد التنمر السيبران ويُنصح بحماية نفسه دائمًا عند القيام .بالأنشطة على وسائل التواصل الاجتماعي، لأنه إذا تُرك دون معالجة، فقد يعيق الأنشطة الإيجابية على وسائل التواصل الاجتماعي ف هذا البحث، ستقوم الباحثة بالاستفادة من التعلم الآلي ف تصنيف نصوص التنمر السيبران المنتشرة ف نصوص التغريدات على منصة تويتر .الطريقة المستخدمة هي شجرة القرار لتصنيف ٤٧,٦٩٢ من بيانات نصوص التغريدات المقسمة إلى ستة أنواع من الفئات وهي :الدين، والجنس، والعمر، والعرق، وتنمر سيبران آخر، وليس تنمرًا سيبرانيًا .بعد المرور بمراحل المعالجة المسبقة للبيانات وهي ث سيتم اختبار البيانات بثلاثة أنواع من سيناريوهات ،TF-IDF تنظيف النصوص وتصفيتها، سيتم حساب وزن البيانات باستخدام تقسيم البيانات، ث تطبيقها على طريقة شجرة القرار بأفضل عمق أقصى بناءً على سيناريوهات البيانات .وأظهرت النتائج المحققة أن سيناريو البيانات ٩٠:١٠ حقق أعلى دقة بنسبة ٨٣,٦٣٪ حيث كان النموذج قادرًا على تصنيف بيانات نصوص التنمر السيبران .بناءً على فئات الجنس، والعمر، والعرق، والدين، بينما كان النموذج أقل كفاءة ف تصنيف فئتي تنمر سيبران آخر وليس تنمرًا سيبرانيًا

BAB I

PENDAHULUAN

1.1 Latar Belakang

Media sosial kini menjadi salah satu sarana komunikasi yang banyak digunakan pada era digital karena jumlah penggunanya terus meningkat. Kemudahan dalam mengakses internet membuat media sosial tidak hanya mempermudah masyarakat dalam berinteraksi, tetapi juga mendukung perkembangan berbagai sektor, terutama dunia bisnis. Di sisi lain, pemanfaatan media sosial juga menimbulkan berbagai dampak negatif. Salah satunya adalah meningkatnya kasus kriminal yang dilakukan melalui media sosial, baik oleh penggunanya secara langsung maupun dengan menjadikan platform tersebut sebagai sarana untuk melakukan tindakan yang melanggar hukum. (Janisar et al., 2021). Perkembangan media sosial saat ini membawa dampak ganda, yakni sebagai sarana informasi sekaligus ruang bagi praktik perundungan daring yang merusak hubungan sosial. Rendahnya kesadaran masyarakat menjadi faktor pendorong utama terjadinya cyberbullying. Berbeda dengan perundungan tradisional yang melibatkan kontak fisik atau secara langsung, cyberbullying menitikberatkan pada penggunaan teknologi komunikasi untuk menyampaikan pesan-pesan ofensif secara berulang. Tindakan ini dilakukan dengan sengaja untuk mendiskreditkan atau merendahkan individu maupun kelompok di ranah digital. (Margono, 2019). Perundungan memberikan dampak multidimensional bagi korbannya, mulai dari gangguan psikologis berupa kecemasan dan depresi, hingga dampak fisik yang

bersifat fatal. Depresi yang timbul akibat perundungan merupakan faktor kritis yang dapat mendorong perilaku bunuh diri. Saat ini, perundungan telah bertransformasi ke dalam bentuk cyberbullying, di mana platform daring dan media sosial digunakan sebagai instrumen untuk melakukan serangan verbal serta merendahkan individu secara sistematis. Hal ini tidak hanya merusak kepercayaan diri korban, tetapi juga mengancam keselamatan jiwa mereka.

Manifestasi cyberbullying dapat ditemukan dalam berbagai bentuk, mulai dari penghinaan fisik (body shaming), diskriminasi rasial, hobi, orientasi seksual, hingga perilaku seksisme. Media sosial menjadi ruang di mana intensitas perundungan siber ini sering kali mencapai titik tertinggi. Twitter, secara spesifik, kerap menjadi medium bagi aksi tersebut melalui unggahan narasi kasar yang bertujuan untuk menyudutkan serta mendelegitimasi reputasi korban. Dampaknya, korban mengalami tekanan psikologis berupa rasa malu dan sakit hati, sementara pelaku memperoleh kepuasan emosional atas keberhasilan tindakannya dalam merusak kredibilitas target (Febriana & Budiarto, 2019). Penyebaran pesan negatif yang cepat di Twitter berkontribusi signifikan terhadap beban emosional korban, yang diperparah oleh faktor anonimitas yang memicu agresivitas pengguna. Kondisi ini menciptakan lingkungan digital yang tidak aman dan mengisolasi korban. Urgensi penanganan cyberbullying terletak pada peningkatan kesadaran publik dan tindakan preventif yang nyata. Mengingat setiap individu berpotensi terlibat sebagai pelaku atau korban, pemahaman mengenai mekanisme perundungan siber menjadi sangat penting. Sifat media sosial yang mampu melipatgandakan jangkauan pesan menjadikan dampak perundungan ini lebih masif

dan berisiko memicu tindakan serupa secara kolektif. Mitigasi terhadap tren ini dapat dicapai melalui edukasi mengenai konten-konten yang berpotensi mengandung unsur perundunganKlasifikasi Upaya mitigasi dampak negatif perundungan siber (cyberbullying) pada platform Twitter menuntut adanya integrasi antara perkembangan teknologi dan prinsip etika Islam. Sebagai agama yang komprehensif, Islam menyediakan kerangka moral bagi umatnya dalam berinteraksi di ruang digital. Dalam konteks ini, nilai-nilai dalam Surah Al-Hujurat ayat 13 mengenai hakikat penciptaan manusia dapat diimplementasikan sebagai landasan teoretis dalam pengembangan sistem klasifikasi tindakan cyberbullying.

’ۡيَاَيُّهَا النَّا س اِنَّا خَلَقۡ ’نكمۡ ۤمنۡ ذَكَ ,,ر وَّاَنۡ ’ثى وَجَعَلۡ ’نكمۡ شعُوۡبًا وَّقَبَاٮِٕلَ لِتَعَارَفـوۡا ۚ اِنَّ اَكۡرَمَكمۡ عِنۡدَ ۡاللّٰ اَتۡ ’قىكمۡ ؕاِنَّ ۡاللّٰ

عَلِيۡ م خَبِيۡ ر

"Wahai manusia, sesungguhnya Kami telah menciptakan kamu dari seorang laki-laki dan perempuan. Kemudian, Kami menjadikan kamu berbangsa-bangsa dan bersuku-suku agar kamu saling mengenal. Sesungguhnya yang paling mulia di antara kamu di sisi Allah adalah orang yang paling bertakwa. Sesungguhnya Allah Maha Mengetahui lagi Maha Teliti." (Q.S Al-Hujurat:13).

Dalam Tafsir Jalalain dijelaskan bahwa Surah Al-Hujurat ayat 13 menerangkan bahwa seluruh manusia diciptakan Allah SWT dari asal-usul yang sama, yaitu Nabi Adam dan Hawa. Selanjutnya, Allah membagi manusia ke dalam berbagai bangsa dan suku sebagai bagian dari keragaman yang telah ditetapkan-Nya. Pada ayat tersebut, istilah *Syu'uuban* merupakan bentuk jamak dari *Sya'bun*, yang menunjukkan tingkatan nasab tertinggi. Di bawahnya terdapat tingkatan suku atau kabilah, kemudian diikuti oleh *Imarah*, *Bathn*, *Fakhdz*, dan tingkatan paling kecil yang disebut *Fashilah*. Sebagai contoh, Khuzaimah termasuk golongan bangsa, Kinanah merupakan kabilah atau suku, Quraisy berada pada tingkatan

Imarah, Qushay termasuk *Bathn*, Hasyim berada pada tingkat *Fakhdz*, sedangkan Al-Abbas merupakan bagian dari *Fashilah*.

Lebih lanjut, frasa *"supaya kalian saling mengenal"* menggunakan kata *Ta'aarafuu*, yang berasal dari bentuk kata *Tata'aarafuu*. Dalam proses pembentukannya, salah satu huruf *Ta* dihilangkan sehingga menjadi *Ta'aarafuu*. Penggunaan kata tersebut menunjukkan bahwa keberagaman bangsa dan suku bukanlah alasan untuk merasa lebih tinggi dari orang lain, melainkan agar manusia dapat saling mengenal, menghargai, dan membangun hubungan yang baik.

Ajaran Islam melalui konsep ta'aarafuu mendekonstruksi paradigma superioritas berbasis nasab dan menggantinya dengan standar ketakwaan sebagai ukuran kemuliaan yang hakiki. Interaksi antarmanusia seharusnya tidak didasari oleh motivasi untuk menyombongkan latar belakang keturunan. Hal ini didasarkan pada keyakinan bahwa Allah memiliki pengetahuan yang komprehensif atas segala kondisi hamba-Nya, termasuk niat-niat yang tersimpan di dalam batin.. Landasan teoretis pengembangan sistem klasifikasi cyberbullying ini merujuk pada prinsip keberagaman identitas manusia yang tertuang dalam Surah Al-Hujurat ayat 13. Berdasarkan Tafsir Jalalain, ayat tersebut menegaskan bahwa kemanusiaan berasal dari asal-usul yang tunggal, yakni Adam dan Hawa. Namun, Allah SWT kemudian membentuk manusia menjadi berbangsa-bangsa (Syu'uuban) dan bersuku-suku (Qaba’il). Dalam tafsir tersebut, dijelaskan adanya hierarki nasab yang sangat terperinci: dimulai dari Syu'uub sebagai tingkatan tertinggi, diikuti oleh suku atau kabilah, kemudian Imarah, Bathn, Fakhdz, hingga unit terkecil yang disebut Fashilah. Struktur ini dapat dicontohkan melalui silsilah dari Khuzaimah

sebagai bangsa, Kinanah sebagai kabilah, Quraisy sebagai Imarah, Qushay sebagai Bathn, Hasyim sebagai Fakhdz, hingga Al-Abbas sebagai Fashilah.

Tujuan utama dari pengelompokan sistematis ini adalah agar manusia dapat saling mengenal (Ta'aarafuu), yang secara linguistik berasal dari kata Tata'aarafuu. Pesan inti dari proses saling mengenal ini adalah untuk membangun harmoni, bukan untuk saling menyombongkan garis keturunan, karena kemuliaan hakiki di sisi Allah hanyalah didasarkan pada tingkat ketakwaan seseorang, bukan pada identitas primordialnya. Allah Maha Mengetahui atas segala kondisi hamba-Nya, termasuk apa yang tersembunyi di dalam batin mereka.

Prinsip teologis mengenai variasi identitas dan pentingnya ta’aruf tersebut menjadi basis dalam merumuskan sistem klasifikasi cyberbullying yang terdiri dari enam kelas: religion, age, gender, ethnicity, other cyberbullying, dan not bullying. Pengkategorian ini mencerminkan kompleksitas identitas manusia yang diuraikan dalam ayat tersebut. Secara spesifik, kelas religion dan ethnicity berfungsi mendeteksi perilaku diskriminatif terhadap keyakinan dan latar belakang etnis, yang selaras dengan konsep bangsa dan kabilah dalam Al-Qur'an. Selanjutnya, kelas age dan gender mencakup perlindungan terhadap variasi identitas individu secara utuh, serupa dengan pembagian sub-suku hingga unit terkecil dalam nasab manusia.

Kategori other cyberbullying disediakan untuk mengidentifikasi tindakan perundungan secara umum di luar atribut identitas spesifik tersebut. Terakhir, kategori not bullying menjadi representasi dari perilaku komunikasi yang ideal sesuai pesan ayat, yakni interaksi yang didasari oleh sikap saling mengenal dan

menghargai tanpa adanya intensi untuk menghakimi atau merendahkan martabat sesama. Dengan mengadopsi sistem klasifikasi yang mencerminkan keragaman identitas manusia ini, penelitian ini berupaya memperkuat nilai kesetaraan dan keadilan dalam memitigasi serta menangani fenomena cyberbullying di ruang digital. Metode klasifikasi merupakan salah satu pendekatan penting dalam upaya penanganan kasus *cyberbullying* (Wang & Potika, 2021). Pendekatan klasifikasi berbasis machine learning memungkinkan identifikasi perilaku perundungan siber secara otomatis, yang pada gilirannya meningkatkan efisiensi penanganan kasus di ruang digital. Di antara berbagai metode klasifikasi, Decision Tree sering menjadi pilihan utama karena kemampuannya dalam menyajikan hasil klasifikasi yang transparan dan mudah dipahami. Keunggulan utama metode ini terletak pada struktur pohon keputusan yang mampu meletakkan pola data secara akurat, menjadikannya instrument yang efektif untuk mengenali karakteristik cyberbullying pada bagian domain dataset (Safavian & Labdgrebe, 1991).

Decision Tree tidak hanya mudah dipahami karena menghasilkan model yang transparan, tetapi juga mampu mengolah data dengan jumlah atribut yang banyak dan mendukung proses klasifikasi ke dalam beberapa kelas. Kemampuan tersebut menjadi salah satu alasan utama metode ini dipilih dalam penelitian klasifikasi teks. Meskipun penelitian mengenai deteksi *cyberbullying* telah banyak dilakukan, penggunaan algoritma Decision Tree untuk kasus tersebut masih belum banyak diterapkan. Kondisi ini menunjukkan bahwa pengembangan metode klasifikasi yang lebih efektif masih diperlukan agar proses identifikasi perilaku *cyberbullying* dapat dilakukan dengan lebih akurat.

Penelitian ini mengkaji penerapan algoritma Decision Tree dalam mengklasifikasikan tindakan cyberbullying, dengan menitikberatkan pada pengembangan model yang akurat sekaligus memiliki tingkat interpretabilitas tinggi. Selain itu, studi ini menganalisis kontribusi tahapan text preprocessing dan ekstraksi fitur Term Frequency-Inverse Document Frequency (TF-IDF) dalam mengoptimalkan performa model. Dengan mengintegrasikan keunggulan Decision Tree dan pemahaman mendalam terhadap pola perundungan siber, penelitian ini berupaya merumuskan strategi deteksi yang komprehensif. Luaran dari studi ini diharapkan dapat berfungsi sebagai instrumen deteksi dini real-time untuk memperkuat ekosistem keamanan pada platform digital.

1.2 Pernyataan Masalah

Bagaimana tingkat performa metode Decision Tree dalam mengklasifikasikan data teks cyberbullying serta kemampuannya dalam mengenali pola klasifikasi yang relevan?

1.3 Tujuan Penelitian

Mengukur dan mengevaluasi performa metode Decision Tree dengan ekstraksi fitur TF-IDF dalam mengklasifikasikan data teks cyberbullying berdasarkan hasil evaluasi model.

1.4 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut.

1. Dataset yang digunakan berasal dari Kaggle dan berupa kumpulan komentar berbahasa Inggris yang diambil dari media sosial Twitter. (https://www.kaggle.com/datasets/andrewmvd/*cyberbullying*-classification).
2. Kelas yang ada pada dataset ini yaitu *gender*, *religion*, *age*, *ethnicity*, *other cyberbullying*, dan *not cyberbullying*.

1.5 Manfaat Penelitian

Penelitian ini mengevaluasi efektivitas algoritma Decision Tree dalam mengidentifikasi cyberbullying dengan menganalisis tingkat akurasi dan hasil klasifikasi yang diperoleh secara mendalam.

BAB II

STUDI PUSTAKA

2.1 Cyberbullying

Cyberbullying merupakan bentuk perilaku agresif yang dilakukan melalui berbagai media digital, seperti media sosial, aplikasi pesan instan, maupun platform komunikasi daring. Tindakan ini dilakukan secara berulang dengan tujuan untuk mengintimidasi, merendahkan, atau menyakiti orang lain. Akibatnya, korban dapat mengalami berbagai dampak negatif, mulai dari gangguan psikologis, emosional, hingga kesulitan dalam kehidupan sosialnya (Kowalski et al., 2019). Menurut Samalo et al. (2024), *cyberbullying* memiliki beberapa bentuk, antara lain penghinaan, penyebaran rumor, pelecehan, dan ujaran kebencian yang disampaikan melalui internet.

Berdasarkan karakteristik tersebut, berbagai penelitian telah mengembangkan metode untuk mendeteksi *cyberbullying* secara otomatis. Salah satunya dilakukan oleh Samalo et al. (2024) yang menerapkan metode Decision Tree untuk mengklasifikasikan *cyberbullying* pada media sosial Twitter. Penelitian tersebut menggunakan dataset berupa *tweet* berbahasa Indonesia yang dibagi ke dalam lima kategori, yaitu *animal bullying*, *sexual harassment*, *appearance bullying*, *general bullying*, dan *non-bullying*. Hasil pengujian menunjukkan bahwa metode Decision Tree mampu memperoleh akurasi sebesar 99,56%, sehingga dinilai efektif dalam mengidentifikasi perilaku *cyberbullying* pada data teks.

Penelitian serupa dilakukan oleh Rahmawati et al. (2023) yang membandingkan kinerja beberapa algoritma klasifikasi dalam mendeteksi *cyberbullying*. Dalam

penelitian tersebut, *Decision Tree* mencapai akurasi sebesar 95,39 %, lebih unggul pada aspek *precision* dan *F1-score* dibandingkan dengan algoritma lain algoritma lain seperti *Logistic Regression* dan KNN. Hal ini menunjukkan bahwa *Decision*

Sementara itu, Noor et al. (2022) melakukan penelitian mengenai deteksi *cyberstalking* dan *cyberbullying* dengan membandingkan algoritma *Naïve Bayes* dan *Decision Tree*. Penelitian ini menggunakan dataset komentar media sosial dengan representasi teks berbasis *bag-of-words*. Hasilnya menunjukkan bahwa *Naïve Bayes* unggul pada nilai *recall* dengan akurasi 95,8 %, sedangkan *Decision Tree* tetap menjadi alternatif yang baik karena kemampuannya menjelaskan proses klasifikasi secara transparan.

Studi yang dilakukan oleh Zaenal (2021) juga meneliti deteksi *cyberbullying* dengan menggunakan metode *Decision Tree*. Penelitian tersebut menggunakan 400 data komentar media yang dibagi ke dalam dua kategori, yaitu *cyberbullying* dan *non-cyberbullying*. Hasil klasifikasi menunjukkan akurasi sebesar 88 %, yang memperlihatkan bahwa *Decision Tree* mampu bekerja cukup baik meskipun dataset yang digunakan relative kecil.

2.2 *Decision Tree*

Decision Tree merupakan salah satu algoritma klasifikasi yang banyak digunakan dalam bidang *machine learning* karena konsepnya sederhana dan hasilnya mudah dipahami. Algoritma ini bekerja dengan membagi data menjadi beberapa kelompok berdasarkan nilai dari atribut tertentu hingga membentuk sebuah struktur pohon keputusan. Proses klasifikasi dimulai dari simpul akar (*root node*) dan berlanjut melalui setiap cabang sesuai dengan kondisi yang telah

ditentukan sampai mencapai simpul akhir (*leaf node*). Setiap simpul di dalam pohon merepresentasikan suatu atribut, sedangkan simpul daun menunjukkan hasil atau kelas yang menjadi keputusan akhir.

Perkembangan metode ini diawali oleh algoritma ID3 (*Iterative Dichotomiser 3*) yang diperkenalkan oleh Quinlan (1986). Selanjutnya, algoritma tersebut terus dikembangkan menjadi C4.5 dan CART (*Classification and Regression Tree*). Decision Tree memiliki beberapa keunggulan, di antaranya mampu mengolah data numerik maupun kategorikal serta menghasilkan model yang mudah diinterpretasikan. Oleh karena itu, algoritma ini sering dimanfaatkan dalam berbagai tugas klasifikasi, termasuk untuk mendeteksi *cyberbullying*.

Dalam konteks tersebut, Decision Tree dinilai sesuai untuk proses klasifikasi teks karena mampu menghasilkan keputusan dengan cepat dan memiliki alur yang transparan. Sementara itu, *cyberbullying* merupakan tindakan agresif yang dilakukan melalui media digital, seperti media sosial, dengan tujuan menyakiti orang lain secara psikologis. Kondisi ini menunjukkan bahwa pengembangan sistem deteksi *cyberbullying* secara otomatis menjadi semakin penting agar interaksi di ruang digital dapat berlangsung dengan lebih aman, terutama bagi kalangan remaja.

Penelitian oleh Pratama dan Wulandari (2021), menggunakan algoritma *Decision Tree* dalam mendeteksi komentar bernuansa *cyberbullying* di *platform* Instagram. Penelitian ini menggunakan fitur ekstraksi teks berupa TF-IDF dan menghasilkan akurasi klasifikasi sebesar 82.67%. Hasil ini menunjukkan bahwa *Decision Tree* cukup efektif dalam menangani teks pendek yang umum ditemukan

dalam komentar media sosial. Selain itu, penelitian yang dilakukan oleh Khairunnisa et al. (2022) membandingkan beberapa algoritma klasifikasi seperti *Naive Bayes*, SVM, dan *Decision Tree* dalam mendeteksi *cyberbullying* pada data *Twitter*. Dalam penelitian tersebut, *Decision Tree* menunjukkan performa yang cukup stabil dengan akurasi sebesar 78.25%, meskipun sedikit lebih rendah dibandingkan SVM. Namun, *Decision Tree* unggul dalam hal interpretabilitas dan visualisasi proses klasifikasi.

Dalam penelitian lainnya oleh Setyawan (2023), algoritma *Decision Tree* dikombinasikan dengan metode pembobotan kata menggunakan *Word2Vec* untuk klasifikasi *cyberbullying* pada komentar *YouTube*. Penelitian ini menunjukkan bahwa integrasi antara representasi vektor kata dan struktur pohon keputusan mampu meningkatkan akurasi klasifikasi hingga 85.10%. Keunggulan dari metode *Decision Tree* adalah kemampuannya dalam memproses data dengan cepat serta mudah diinterpretasikan, yang sangat berguna dalam proses moderasi konten secara otomatis. Namun, kekurangannya adalah kecenderungan terhadap *overfitting* jika struktur pohon terlalu kompleks. Oleh karena itu, pemangkasan pohon (*pruning*) menjadi teknik penting dalam implementasi algoritma ini.

Dari beberapa penelitian tersebut, dapat disimpulkan bahwa metode *Decision Tree* memiliki potensi besar dalam deteksi dan klasifikasi teks *cyberbullying*. Kombinasi dengan teknik praproses dan pembobotan fitur yang tepat dapat meningkatkan performa sistem secara signifikan.

Tabel 2.1 Penelitian Terkait

No	Nama	Judul Penelitian	Kesamaan	Perbedaan

1.	Samalo et al.(2024)	Perbandingan Algoritma Klasifikasi dalam Mendeteksi Cyberbullying pada Data Teks	- Membahas tentang *cyberbullying* pada media sosial - Data berasal dari media sosial *twitter* - Menggunakan metode klasifikasi yang sama	- Lebih berorientasi peforma model *Decision Tree* dalam klasifikasi *multi class cyberbullying*
2.	Rahmawati et al. (2023)	Analisis Perbandingan Algoritma Klasifikasi dalam Mendeteksi *Cyberbullying* pada Data Teks	- Membahas tentang *cyberbullying* pada media sosial - Menggunakan algoritma *Decision tree* dalam penelitiannya	- Menggunakan perbandingan metode *Decision* tree dengan *Logistic Regression* dan KKN - Menjelaskan keunggulan *Decision tree* dibanding metode lain
3.	Noor et al. (2022)	Pendekatan Pembelajaran Mesin dalam Deteksi *Cyberstalking* pada Media Sosial Menggunakan *Naive Bayes* dan *Decision Tree*	- Membahas tentang *cyberbullying* pada media social	- Menggunakan ekstraksi fitur *Bag-of-Words* - Menekankan perbandingan *recall* dan akurasi - Menggunakan 2 metode dalam penelitiannya yaitu *Naïve Bayes* dan *Decision Tree*
No.	**Nama**	**Judul Penelitian**	**Kesamaan**	**Perbedaan**

4.	Zaenal (2021)	Deteksi Cyberbullying pada Komentar Media Sosial Menggunakan Algoritma *Decision Tree*.	- Membahas tentang *cyberbullying* pada media sosial - Data berasal dari media sosial *twitter* - Menggunakan algoritma *Decision tree*	- Hanya menghasilkan dua *class* - Jumlah *dataset* terbatas hanya 400 data komentar
6.	Khairunnisa et al. (2022)	Perbandingan Algoritma Klasifikasi *Naive Bayes*, SVM, dan *Decision Tree* untuk Deteksi *Cyberbullying* pada Data *Twitter*	- Membahas tentang *cyberbullying* pada media sosial - Data berasal dari media sosial twitter	- Membandingkan algoritma *Naïve Bayes*, SVM, dan *Decision tree*
7.	Setyawan (2023)	Klasifikasi Cyberbullying pada Komentar YouTube dengan Menggunakan Word2Vec dan Decision Tree	- Membahas tentang *cyberbullying* Menggunakan algoritma Decision tree	- Menggunakan esktraksi fitur *Word2Vec* - Menggunakan dataset media sosial *Youtube*

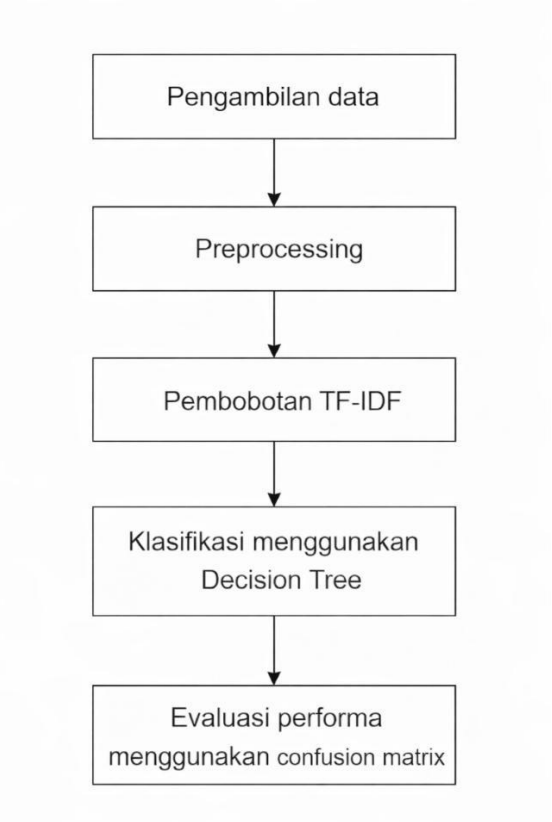
Berdasarkan pemaparan pada Tabel 2.1, dapat diketahui bahwa penelitian ini memiliki relevansi yang kuat dengan beberapa studi terdahulu, terutama pada objek penelitian yaitu klasifikasi *cyberbullying* pada media sosial. Namun, perbedaan mendasar terletak pada penggunaan algoritma *Decision tree* yang dikombinasikan dengan pembobotan TF-IDF untuk mencapai tingkat akurasi yang lebih optimal.

BAB III

DESAIN DAN IMPLEMENTASI

3.1 Desain Penelitian

Desain penelitian ini menjadi landasan operasional yang mengarahkan setiap fase kegiatan dalam proses penelitian secara terstruktur. ini ditampilkan dalam bentuk diagram pada Gambar 3.1



Gambar 3.1 Desain Penelitian

Alur kerja sistem ini diawali dengan akuisisi dataset dari platform Kaggle, yang selanjutnya dipartisi menjadi data latih (training data) dan data uji (testing data). Dataset tersebut mencakup teks tweet multibahasa (Inggris dan Indonesia) yang terdistribusi ke dalam enam label kelas, yaitu: not_cyberbullying, gender, religion, age, ethnicity, dan other_cyberbullying. Sebelum memasuki tahap

pemodelan, data melalui proses preprocessing guna memastikan kesiapan data untuk dianalisis. Tahap berikutnya melibatkan penerapan metode pembobotan Term Frequency-Inverse Document Frequency (TF-IDF) pada data latih untuk membangun model klasifikasi. Terakhir, algoritma Decision Tree diimplementasikan untuk melakukan klasifikasi, yang kemudian kinerjanya dievaluasi menggunakan Confusion Matrix.

3.2 Pengumpulan Data

Pada penelitian ini pengumpulan data dan informasi menggunakan data sekunder *cyberbullying* komentar berbahasa inggris yang didapat dari *Kaggle* dengan dataset yang digunakan *Cyberbullying Classification* yaitu dengan alamat (https://www.kaggle.com/datasets/andrewmvd/*cyberbullying*-classification). Pada penelitian ini menggunakan keseluruhan data yang telah disediakan pada *Kaggle* dengan pembagian data kedalam enam kelas. Data tersebut mencakup komentar yang diklasifikasikan ke dalam enam kategori, yaitu *age*, *ethnicity*, *gender*, *religion*, *other cyberbullying*, dan *not cyberbullying* contoh sampel data dari masing masing kelas dapat dilihat pada Tabel 3.1.

Tabel 3.1 Pengumpulan Data

Dokumen	Komentar	Kelas
D1	*@BelieveTed i feel so fucking bad for you. You dumb fucking nigger im Doxing you and im glad you live in michigan im beating the fuck outa u*	*ethnicity*
D2	*Nobody asks to be made fun of but it happens to everyone gay jokes straight jokes rape jokes all of its gonna happen*	*gender*
D3	*@biebervalue @greenlinerzjm You fucking moron, slavery is approved right in the Quran. Read it you stupid idiot*	*religion*
D4	*Just realized that all the girls who used to bully me in middle school and hs are either single moms, divorced, or dating someone who cheats on them and if thatâ€™s not the best karma ever idk what is*	*age*

D5	*That was fun. Now it's over. This is why I can't even consider sites like TechRaptr as anything more than hit piece blogs.*	*not_cyberbullying*
D6	*Both of the dogs are staring at me because they didn't know I could make noises this high pitched.*	*other_cyberbullying*

Jumlah data pada setiap sentimen, ditampilkan pada Tabel 3.2.

Tabel 3.2 Jumlah Data Awal

Sentimen	**Jumlah Data Awal**
Religion	7.998
Gender	7.992
Ethnicity	7.973
Age	7.961
Other Cyberbullying	7.945
Not Cyberbullying	7.823
Total	**47.692**

Berdasarkan Tabel 3.2, diketahui bahwa data yang digunakan total sebanyak 47.692 data, dimana untuk setiap sentimen diketahui jumlah datanya tidak sama banyak, hal ini menunjukkan bahwa data yang diambil pada *Kaggle* merupakan data yang tidak seimbang (*imbalance*), sehingga nantinya saat penelitian akan digunakan teknik atau cara untuk menyeimbangkan jumlah data pada masing-masing kelas, untuk membantu model Decision Tree agar mampu mengklasifikasikan data teks *cyberbullying*. Untuk mengetahui angka pasti berikut akan ditampilkan Tabel 3.3 yang menunjukkan jumlah hashtag, simbol, mention, hyperlink, emoji dan spasi berlebih.

Tabel 3.3 Jumlah *Special Character* pada Data

Sentimen	**Jumlah**				
	Hashtag	***Mention***	***Hyperlink***	***Emoji***	**Spasi Berlebih**
Religion	816	2.152	498	9.062	269

Age	391	348	159	8.876	3
Gender	1.519	2.867	601	8.674	291
Ethnicity	651	3.205	316	9.067	10
Other Cyberbullying	666	2.754	708	5.080	186
Not Cyberbullying	2.143	3.210	811	7.277	431
Total	**6.186**	**14.536**	**3.093**	**48.036**	**1.190**

Pada tahap normalisasi teks, dilakukan pembersihan *hashtag*, *mention*, hyperlink, emoji, dan spasi berlebih sehingga nantinya data tweet hanya berisi teks saja, hal ini untuk membuat model mampu mengolah data teks yang akan digunakan, Tabel 3.3 menunjukkan jumlah simbol, spasi berlebih, mention, hyperlink, dan hashtag pada setiap sentimen di mana masing-masing akan dihapus kemudian untuk teks yang memilki huruf kapital juga akan diubah menjadi huruf kecil.

Tabel 3.4 Jumlah Data Duplikat

Sentimen	**Jumlah Data Duplikat**
Religion	48
Gender	107
Ethnicity	299
Age	204
Other Cyberbullying	287
Not Cyberbullying	308
Total	**1.253**

Tabel 3.4 menunjukkan jumlah data duplikat yang terdistribusi pada setiap kelas sentimen, jumlah data duplikat ini menunjukkan bahwa pada setiap sentimen masih terdapat data yang sama sehingga nantinya akan dihapus pada tahap *cleaning*.

Tabel 3.5 Minimal dan Maksimal Jumlah Kata

Sentimen	**Jumlah Data**	
	Kurang Dari 3 Kata	**Lebih Dari 100 Kata**
Religion	16	0
Gender	20	1
Ethnicity	257	3
Age	17	3
Other Cyberbullying	1.301	3
Not Cyberbullying	914	2
Total	**2.525**	**12**

Tabel 3.5 menunjukkan jumlah data maksimal dan minimal yang terdistribusi pada setiap kelas sentimen, analisis jumlah data ini nanti bertujuan untuk menyaring jumlah kata yang terlalu banyak seperti kurang dari 3 kata karena kurangnya konteks dan makna ,serta lebih dari 100 kata dengan tujuan efisiensi dan kinerja model. Jumlah data yang digunakan dalam penelitian mengalami pengurangan di mana pada data setelah tahapan *preprocessing* data dielminasi untuk memudahkan model dalam memperoleh performa dan memahami data teks tweet yang akan diolah dalam fitur pembobotan dan metode *Decision Tree.* Analisis jumlah data sebelum dan sesudah *preprosesing* dapat dilihat pada tabel 3.11

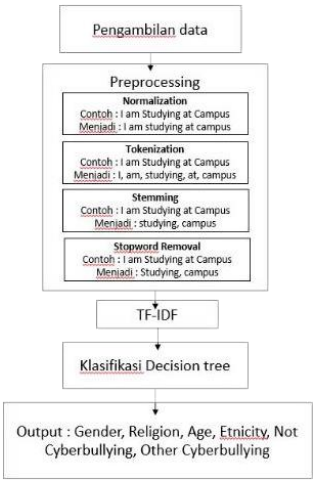
Tabel 3.6 Analisis Jumlah Data Sebelum dan Sesudah Preproses

Sentimen	**Jumlah Sebelum *Preprosessing***	**Jumlah Setelah *Preprosessing***
Religion	7998	7893
Age	7992	7830
Gender	7973	7311
Ethnicity	7961	7707
Not Cyberbullying	7945	6371
Other Cyberbullying	7823	4669
Total	47692	41781

3.3 Desain Sistem

Penelitian ini memiliki sebuah sistem klasifikasi yang dirancang untuk menggambarkan tahap-tahap yang digunakan pada sistem, dimulai dari

pengambilan data hingga menghasilkan data yang telah diklasifikasikan. Setelah itu, sistem akan mengevaluasi kinerjanya dengan menggunakan *confusion matrix*. Rancangan desain sistem yang akan diimplementasikan dalam penelitian ini dapat dilihat pada Gambar 3.2



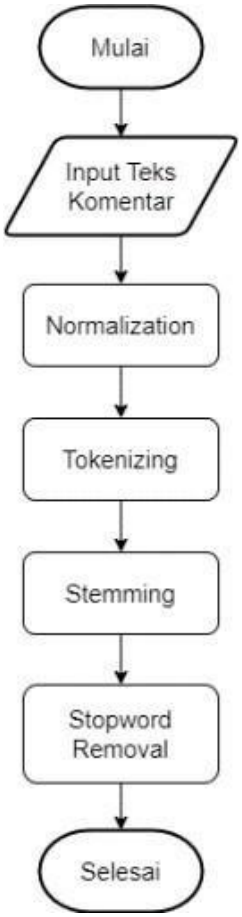
Gambar 3.2 Desain Sistem

Dalam penelitian ini, sistem klasifikasi yang dirancang memiliki proses awal dengan mengambil data dari sumber *Kaggle* yang relevan. Setelah data diambil, dataset ini menjalani serangkaian tahapan penting yang dimulai dengan *preprocessing* data untuk memastikan kebersihan dan kualitasnya, termasuk penghapusan data yang tidak relevan dan normalisasi teks. Selanjutnya, dilakukan pembobotan dengan metode TF-IDF untuk menghasilkan representasi numerik dari teks. Algoritma klasifikasi *Decision Tree* digunakan untuk melakukan klasifikasi sentimen pada data tersebut. Hasil klasifikasi akan menjadi enam kelas yaitu "*gender*", "*religion*", "*age*", "*ethnicity*", "*other_cyberbullying",* dan "*not_cyberbullying*". Untuk mengukur kinerja sistem, sistem evaluasi dilakukan dengan *confusion matrix* untuk mengukur *accuracy*, *precision*, *recall*, *dan F1-*

score. Dengan demikian, sistem ini tidak hanya mengklasifikasikan data, tetapi juga melibatkan serangkaian langkah yang komprehensif dalam pengolahan dan evaluasi data untuk menghasilkan hasil yang akurat.

3.4 *Preprocessing*

Preprocessing merupakan langkah penting dalam mengolah teks untuk membuatnya lebih terstruktur dan lebih mudah dipahami oleh sistem, sehingga memfasilitasi proses klasifikasi. Dalam tahap ini, teks mengalami serangkaian proses untuk mempersiapkannya secara optimal. Proses normalisasi teks, termasuk mengubah huruf menjadi huruf kecil dan menghapus tanda baca serta gambar ekspresi, hal ini dilakukan untuk memastikan konsistensi data. Kemudian tokenisasi, yang memecah teks menjadi potongan kata yang disebut token. *Stemming* diterapkan untuk mengubah kata-kata ke bentuk dasarnya, sehingga variasi kata dengan makna yang sama dapat diidentifikasi dengan benar. Selanjutnya, dilakukan penghapusan *stop words*, yaitu kata-kata dangkal yang tidak memiliki makna signifikan dalam analisis. Langkah terakhir teks diubah menjadi representasi numerik menggunakan teknik seperti TF- IDF Tahapan pada *preprocessing* dapat dilihat pada Gambar 3.3



Gambar 3.3 *Flowchart Preprocessing*

3.4.1 *Normalization*

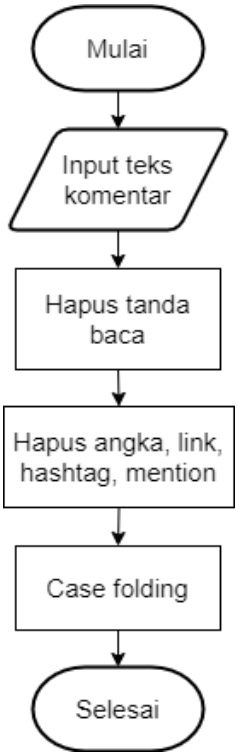
Normalization digunakan untuk menghapus perbedaan-perbedaan umum yang terdapat pada kumpulan data teks opini. Tahapan yang dilakukan dalam *normalization* meliputi penghapusan angka, penghapusan spasi kosong, penghapusan tanda hubung, penghapusan tanda baca, penghapusan gambar karakter, penghapusan *hyperlink*, dan *case folding*, yaitu mengubah semua huruf besar menjadi huruf kecil. Tujuan dari normalisasi ini adalah untuk menyederhanakan teks dan membuatnya lebih konsisten sehingga memudahkan proses analisis, klasifikasi, dan ekstraksi informasi dari kumpulan data teks opini. Dengan melakukan normalisasi, dapat memastikan bahwa data teks yang diolah

lebih mudah diproses secara komputasi dan memberikan hasil yang lebih akurat. Tahap *normalization* dapat dilihat pada Tabel 3.3.

Tabel 3.7 Contoh Normalisasi

Sebelum *Normalization*	*@BelieveTed i feel so fucking bad for you. You dumb fucking nigger im Doxing you and im glad you live in michigan im beating the fuck outa u*
Setelah *Normalization*	*i feel so fucking bad for you you dumb fucking nigger im doxing you and im glad you live in michigan im beating the fuck outa u*

Diagram alur proses normalization dapat dilihat pada Gambar 3.4 berikut.



Gambar 3.4 *Flowchart Normalization*

3.4.2 *Tokenizing*

Tahap *tokenizing* adalah tahapan teks atau kalimat komentar diuraikan menjadi kata yang lebih kecil, yang disebut token atau term (Aninditya et al., 2019). *Tokenizing* ini dapat dianggap sebagai bentuk segmentasi teks. Proses ini penting untuk memudahkan analisis dan pemrosesan lebih lanjut dari teks. Dalam analisis

teks, *tokenizing* berfungsi sebagai langkah awal untuk ekstraksi informasi dan pembuatan model teks. Tahap *Tokenizing* dapat dilihat pada Table 3.5

Tabel 3.8 Contoh Tokenizing

Sebelum *Tokenizing*	*[i feel so fucking bad for you you dumb fucking nigger im doxing you and im glad you live in michigan im beating the fuck outa u]*
Setelah *Tokenizing*	*["i", "feel", "so", "fucking", "bad", " for", " you", " you", "dumb", "fucking","nigger", "im", "doxing", " you", "im", "glad", " you", "live", "in", "michigan", "im", "beating", "the", "fuck", "outa", "u"]*

3.4.3 *Stemming*

Pada tahap *stemming*, kata diubah menjadi bentuk dasarnya yang disesuaikan dengan struktur bahasa yang dipergunakan (Aninditya et al., 2019). Langkah yang dilakukan pada tahap ini adalah menghapus imbuhan yang terdapat pada kata untuk mendapatkan bentuk kata aslinya. Proses *stemming* dalam penelitian ini menggunakan *library Python* yaitu *Natural Language Tool Kit* (NLTK) yang dapat diakses melalui tautan berikut https://www.nltk.org/. Tahap *Stremming* dapat dilihat pada Table 3.8.

Tabel 3.9 Contoh *Stemming*

Sebelum *Stemming*	["i", "feel", "so", "fucking", "bad", " for", " you", " you", "dumb", "fucking","nigger", "im", "doxing", " you", "im", "glad", " you", "live", "in", "michigan", "im", "beating", "the", "fuck", "outa", "u"]
Setelah *Stemming*	["i", "feel", "so", "fuck", "bad", " for", " you", " you", "dumb", "fuck","nigger", "im", "dox", " you", "im", "glad", " you", "live", "in", "michigan", "im", "beat", "the", "fuck", "outa", "u"]

3.4.4 *Stopword Removal*

Stopword removal merupakan sebuah proses dimana kata-kata umum atau yang disebut dengan *stopword* akan dihilangkan dari sebuah teks karena kata-kata tersebut tidak memberikan kontribusi penting dalam analisis teks (Sarica & Luo, 2021). Contohnya, kata-kata seperti kata depan, kata sambung, dan kata-kata umum

lainnya. Proses *stopword removal* dilakukan agar proses analisis teks dapat berjalan lebih cepat dan juga dapat mengurangi dimensi data yang tidak perlu, sehingga dapat meningkatkan kualitas dari analisis yang dilakukan. *Stopword removal* merupakan sebuah proses penting dalam pengolahan teks untuk menghilangkan kata-kata umum yang tidak memberikan kontribusi signifikan dalam analisis teks. Hal ini membantu mempercepat proses analisis dan mengurangi dimensi data yang tidak relevan, yang pada gilirannya meningkatkan efisiensi dan akurasi dari analisis teks yang dilakukan. Contoh *Stopward removal* dapat dilihat pada Tabel 3.9.

Tabel 3.10 Contoh *Stopword removal*

Sebelum *Stopword removal*	["i", "feel", "so", "fuck", "bad", " for", " you", " you", "dumb", "fuck","nigger", "im", "dox", " you", "im", "glad", " you", "live", "in", "michigan", "im", "beat", "the", "fuck", "outa", "u"]
Setelah *Stopword removal*	["feel", "fuck", "bad", "dumb", "fuck","nigger", "dox", "glad", "live", "michigan", "beat", "fuck"]

3.4.5 Data *Filtering* dan *Cleaning*

Tahap penyaringan dan pembersihan kata, di mana pembersihan dilakukan dengan cara menghapus data yang sama persis (duplikasi) untuk menjaga kualitas data, kemudian penyaringan data dengan membatasi jumlah kata dalam setiap teks tweet yakni untuk kalimat yang terlalu pendek dan tidak memiliki konteks yang cukup untuk diklasifikasikan ke dalam teks *cyberbullying*, serta kalimat yang terlalu panjang yang cenderung mengandung teks yang tidak relevan dan justru menggangu performa dari *Decision Tree.* Tabel untuk menunjukkan data setelah penyaringan dan pembersihan dapat dilihat pada tabel 3.10.

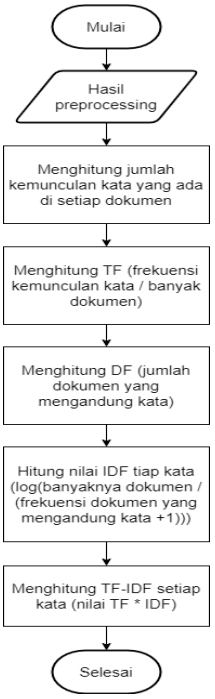
Tabel 3.11 Data Setelah *Filtering* dan *Cleaning*

Teks	Jumlah Kata
sweetieniesha im get bulli via twitter thevics...	92
andreagcav viviaanajim recuerda como nosotra t...	85
cat look like frog would realiz nasti cruel li..	80

3.5 TF-IDF

Dalam penelitian ini, penulis telah memilih metode *Term Frequency- Inverse Document Frequency* (TF-IDF) sebagai strategi utama untuk memberikan bobot pada kata-kata dalam dataset yang digunakan. Proses dimulai dengan tahap *preprocessing*, di mana data disiapkan dan dibersihkan untuk analisis lebih lanjut. Setelah tahap ini selesai, kata-kata dalam dataset menjadi term yang akan menjadi fokus dalam langkah-langkah berikutnya.. Algoritma ini memberikan bobot yang lebih besar kepada kata-kata yang jarang muncul dalam dataset, sehingga memiliki nilai informatif yang lebih tinggi (Imamah & Rachman, 2020). *Term Frequency* (TF) merupakan ukuran tingkat keseringan sebuah kata muncul pada sebuah teks. Dengan kata lain, ini adalah ukuran relatif dari frekuensi kemunculan sebuah kata dalam sebuah dokumen. Semakin sering sebuah kata muncul, semakin tinggi pula nilai TF-nya untuk dokumen tersebut. *Inverse Document Frequency* (IDF) merupakan sebuah algoritma yang dipergunakan dalam menghitung probabilitas kebalikan dari kemunculan sebuah kata dalam seluruh teks. IDF memperhitungkan seberapa umum atau jarang suatu kata dalam kumpulan dokumen. Kata-kata yang sering muncul dalam banyak dokumen akan memiliki nilai IDF yang rendah, sedangkan kata-kata yang jarang muncul dalam dokumen akan memiliki nilai IDF yang tinggi. Hal ini memungkinkan untuk memberikan bobot yang lebih tinggi pada kata-kata yang jarang muncul namun memiliki kepentingan yang besar dalam suatu konteks. Dengan menggabungkan TF dan IDF, metode TF-IDF memberikan bobot yang lebih baik pada kata-kata dalam dataset dengan memperhitungkan kedua faktor tersebut. Hal ini membantu dalam menyoroti kata-kata yang memiliki

relevansi dan signifikansi tertentu dalam suatu analisis, seringkali meningkatkan kemampuan model atau algoritma dalam memahami dan mengekstraksi informasi yang relevan dari dataset yang dianalisis. Ilustrasi dari alur pada proses perhitungan TF-IDF dapat dilihat pada Gambar 3.5



Gambar 3.5 *Flowchart* TF-IDF

Pada tabel 3.10 menunjukkan jumlah kemunculan kata pada setiap dokumen setelah melalui tahap *preprocessing*. Setiap nilai 1 menunjukkan bahwa kata tersebut muncul satu kali pada dokumen terkait, sedangkan nilai 0 menunjukkan kata tidak muncul. Data ini selanjutnya digunakan sebagai dasar pembentukan bobot TF sebelum dilakukan pembobotan TF-IDF dan proses klasifikasi menggunakan algoritma *Decision Tree*.

Tabel 3.12 Jumlah kata yang muncul dalam kelas

Kata	**Jumlah Kata**					
	D1	**D2**	**D3**	**D4**	**D5**	**D6**
Fuck	1	0	0	0	0	1

Call	1	0	0	1	0	0
Feel	1	0	0	0	0	0
Bad	1	0	0	0	0	1
Dumb	1	0	0	0	0	0
Nigger	1	0	0	0	0	1
Dox	1	0	0	0	0	0
Glad	1	0	0	0	0	0
Live	1	0	0	0	0	0
michigan	1	0	0	0	0	0
Beat	1	0	0	0	0	1
Bully	0	1	0	0	0	0
Rude	0	1	0	0	0	0
School	0	1	0	0	0	0
Destroy	0	0	1	0	0	0
Life	0	0	1	0	0	1
Welcom	0	0	1	0	0	0
World	0	0	1	0	0	0
Islam	0	0	1	0	0	0
Terror	0	0	1	0	0	0
Female	0	0	0	1	0	0
Basic	0	0	0	1	0	0
Bitch	0	0	0	1	0	0
Start	0	0	0	0	1	0
thursday	0	0	0	0	1	0
Class	0	0	0	0	1	0
Cancel	0	0	0	0	1	0
tomorrow	0	0	0	0	1	0
Total	10	3	6	3	5	5

Kemudian dihitung menggunakan Algoritma TF-IDF sesuai persamaan 3.1. Hal pertama yang dilakukan adalah menghitung *Term Frequency* (TF).

$$TF_{ij} = \frac{f_{ij}}{max\, k f_{kj}} \quad (3.1)$$

Dalam persamaan 3.1, f_{ij} merujuk pada frekuensi istilah i dalam dokumen j, sementara $max_k f_{kj}$ mengacu pada frekuensi fitur paling umum atau fitur dengan frekuensi tertinggi dalam dokumen tersebut. Perhitungan *Term Frequency* dapat

dilihat pada Table 3.11.

Tabel 3.13 Perhitungan *Term Frequency*

Dokumen	Jumlah Dokumen	Kata	Frekuensi	TF
D1	10	call	1	1/11
D2	3		0	0/3
D3	6		0	0/6
D4	3		1	1/3
D5	5		0	0/5
D6	5		0	0/5

Selanjutnya dilakukan perhitungan Inverse Document Frequency untuk melakukan keseimbangan antara distribusi nilai. Ini dilakukan dengan rumus berikut:

$$IDF = log\ (\frac{N}{n_i}) + 1 \tag{3.2}$$

Dalam persamaan 3.2, N merupakan jumlah total dokumen pelatihan yang digunakan dan n_i adalah adalah jumlah dokumen latih yang mengandung istilah i. Perhitungan IDF dapat dilihat pada Tabel 3.12.

Tabel 3.14 Perhitungan IDF

Dokumen	Frekuensi	Kata	IDF
D1	1	call	$log\ \frac{5}{2} + 1 = 1.39794$
D2	0		
D3	0		
D4	1		
D5	0		
D6	0		

TF-IDF dihitung sesuai dengan persamaan 3.3

$$TFIDF_{ij} = TF_{ij} \times IDF_i \tag{3.3}$$

Persamaan 3.3 menggabungkan nilai Term Frequency (Tf_{ij}) dan Inverse Document Frequency (IDF_i) untuk menghasilkan skor bobot untuk istilah *i* dalam dokumen *j*. Bobot TF-IDF digunakan menilai seberapa signifikan sebuah kata dalam sebuah dokumen relatif terhadap keseluruhan koleksi dokumen. Nilai TF mengukur seberapa sering sebuah kata muncul dalam sebuah dokumen, sementara IDF mengukur seberapa unik atau jarang kata tersebut muncul dalam keseluruhan korpus dokumen. Perhitungan TF-IDF dapat dilihat pada Tabel 3.9.

Tabel 3.15 Perhitungan TF-IDF

Dokumen	Kata	TF	IDF	TF-IDF
D1	*call*	1/11	1.39794	0.12708
D2		0/3		0
D3		0/6		0
D4		1/3		0.46598
D5		0/5		0
D6		0/5		0

Dengan mengalikan TF dan IDF, kita mendapatkan bobot yang lebih tinggi untuk kata-kata yang muncul secara sering dalam sebuah dokumen tetapi jarang muncul dalam dokumen lainnya, dan sebaliknya. Proses ini membantu melihat kata-kata yang paling relevan dan signifikan dalam sebuah dokumen dalam konteks klasifikasi atau analisis teks. Selanjutnya, nilai TF-IDF ini dapat digunakan sebagai fitur untuk melatih model pembelajaran mesin atau digunakan klasifikasi dokumen. Penggunakan nilai TF-IDF sebagai fitur, model dapat dilatih untuk melakukan klasifikasi dokumen dengan lebih akurat dan efisien. Berikut ini merupakan beberapa hasil perhitungan TF-IDF dalam dokumen dapat dilihat pada Tabel 3.14

*T*abel 3.16 Hasil Perhitungan TF-IDF

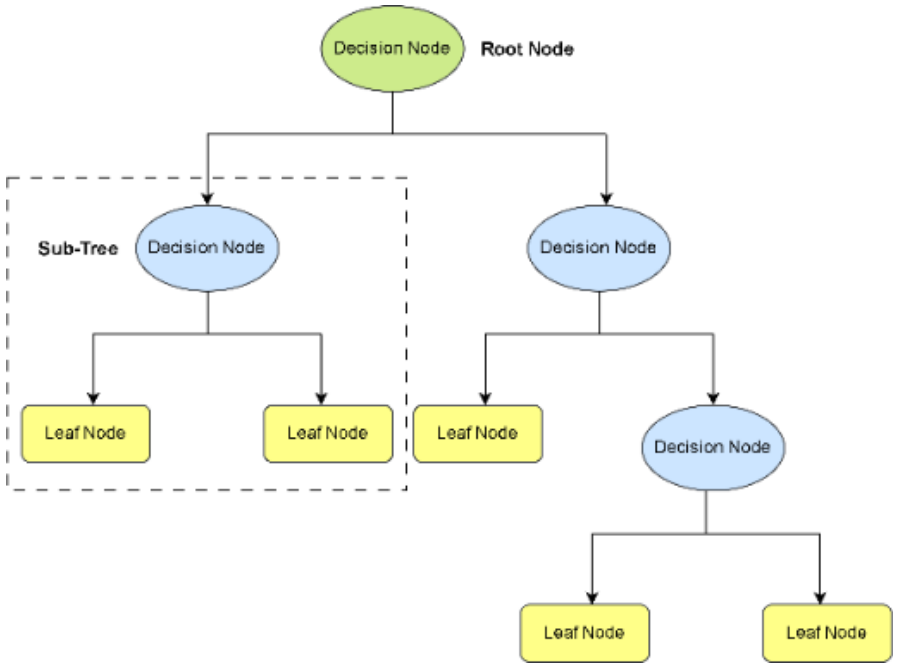
Kata	Jumlah Kata					
	D1	D2	D3	D4	D5	D6
fuck	0.15444	0	0	0	0	0.42181
call	0.12708	0	0	0.46598	0	0

feel	0.15444	0	0	0	0	0
bad	0.15444	0	0	0	0	1
dumb	0.15444	0	0	0	0	0
nigger	0.15444	0	0	0	0	0.32512
dox	0.15444	0	0	0	0	0
glad	0.15444	0	0	0	0	0
live	0.15444	0	0	0	0	0
michigan	0.15444	0	0	0	0	0
beat	0.15444	0	0	0	0	0.25122
bully	0	0.5663	0	0	0	0
rude	0	0.5663	0	0	0	0
school	0	0.5663	0	0	0	0
destroy	0	0	0.28315	0	0	0
life	0	0	0.28315	0	0	0.23512
welcom	0	0	0.28315	0	0	0
world	0	0	0.28315	0	0	0
islam	0	0	0.28315	0	0	0
terror	0	0	0.28315	0	0	0
female	0	0	0	0.42472	0	0
basic	0	0	0	0.42472	0	0
bitch	0	0	0	0.42472	0	0
start	0	0	0	0	0.33979	0
thursday	0	0	0	0	0.33979	0
class	0	0	0	0	0.33979	0
cancel	0	0	0	0	0.33979	0
tomorrow	0	0	0	0	0.33979	0

3.6 Decision Tree

Metode *Decision Tree* adalah salah satu algoritma *supervised learning* yang bekerja dengan membangun struktur pohon keputusan berdasarkan atribut-atribut yang terdapat pada data. Metode ini dapat digunakan untuk klasifikasi maupun regresi, namun pada penelitian ini *Decision Tree* difokuskan untuk klasifikasi teks cyberbullying**.** Algoritma ini bekerja dengan melakukan pembagian (*splitting*) data secara berulang menggunakan ukuran tertentu—seperti Gini Impurity atau *Entropy*—hingga diperoleh struktur pohon yang mampu memisahkan kelas dengan sebaik mungkin (Han et al., 2012).

Arsitektur metode *Decision Tree* pada penelitian ini dapat dilihat pada Gambar 3.6, yang menggambarkan proses pembentukan pohon mulai dari *node* akar hingga *node* daun. Arsitektur metode Decision Tree pada penelitian ini dapat dilihat pada Gambar 3.6, yang menggambarkan proses pembentukan pohon mulai dari node akar hingga node daun.

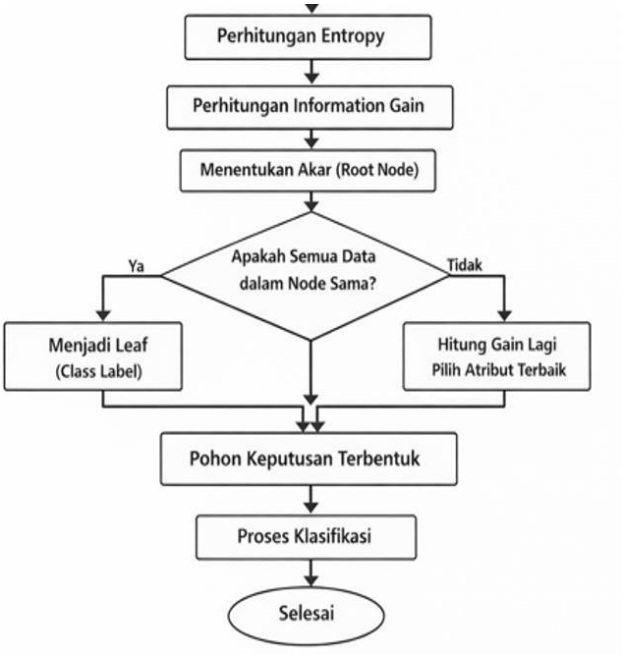


Gambar 3.6 Diagram *Decision Tree*

Pada Gambar 3.6 terlihat bahwa *Decision* Tree terdiri atas beberapa komponen utama, yaitu *node* akar (*root node*)**,** node internal**,** dan *node* daun (*leaf node*)**.** *Node* akar adalah titik awal proses klasifikasi di mana atribut terbaik dipilih melalui perhitungan informasi seperti *Gini Index* atau *Information Gain*. Setelah atribut tersebut dipilih, dataset dibagi menjadi beberapa subset berdasarkan nilai atribut tersebut. Pembagian ini terus dilakukan secara rekursif pada *node-node*

internal sampai tidak ada lagi atribut yang dapat digunakan atau kondisi *stopping criteria* terpenuhi. Hasil akhir dari proses pembagian ini adalah node daun yang berisi kelas akhir (F), yaitu kategori yang telah diprediksi model.

Dalam konteks penelitian ini, *node-node* pada *Decision Tree* bekerja dengan menggunakan fitur berbasis TF-IDF, sehingga setiap input berupa vektor numerik yang merepresentasikan bobot kemunculan kata pada teks. Vektor inilah (X) yang menjadi dasar bagi pohon keputusan untuk menentukan jalur keputusan yang paling tepat hingga mencapai kelas tertentu. Untuk *flowchart Decision Tree* dapat dilihat pada Gambar 3.7.



Gambar 3.7 *Flowchart Decision Tree*

Berikut ini merupakan Tahapan Algoritma serta contoh perhitungan dari metode *Decision tree :*

a. Dataset disusun dari fitur hasil TF-IDF dan label kelas. Data ini digunakan sebagai data latih untuk membangun model.

b. Selanjutnya melakukan perhitungan Entropy awal, Entropy digunakan untuk mengukur tingkat ketidakpastian data berdasarkan distribusi kelas. Nilai entropy menjadi dasar dalam menentukan kualitas pemisahan data dapat dilihat pada Persamaan 3.4.

$$Entropy\ (S_i) = -\sum p_{ij}\ log_2\ (p_{ij}) \quad (3.4)$$

c. Kemudian melakukan entropy setiap fitur, Setiap fitur teks digunakan untuk membagi data, kemudian dihitung entropy pada masing-masing hasil pembagian tersebut untuk melihat seberapa baik fitur memisahkan kelas. Perhitungan berdasarkan Persamaan 3.5.

$$Entropy\ (S_i) = -\sum p_{ij}\ log_2\ (p_{ij}) \quad (3.5)$$

d. Melakukan $SplitInfo(A) = -\sum \frac{|S_j|}{|S|} \log_2 \frac{|S_j|}{|S|}$ perhitungan Information Gain yang menunjukkan seberapa besar pengurangan entropy setelah data dipisahkan oleh suatu fitur. Semakin besar nilainya, semakin baik fitur tersebut. Perhitungan berdasarkan Persamaan 3.6.

$$Gain(S, A) = Entropy(S) - \sum \frac{|S_j|}{|S|} Entropy(S_j) \quad (3.6)$$

e. Melakukan perhitungan Split Information yang digunakan untuk menilai kualitas pembagian data oleh suatu fitur agar tidak terjadi bias pada fitur tertentu. Perhitungan berdasarkan Persamaan 3.7.

(3.7)

f. Melakukan perhitungan Gain Ratio yang merupakan hasil pembagian antara Information Gain dan Split Information. Fitur dengan Gain Ratio tertinggi dipilih sebagai node pohon. Perhitungan berdasarkan Persamaan 3.8.

$$GainRatio(A) = \frac{Gain(S, A)}{SplitInfo(A)} \quad (3.8)$$

g. Penentuan Node dan Pembentukan PohonFitur terbaik dijadikan node, lalu data dibagi ke dalam cabang-cabang. Proses ini diulang hingga terbentuk node daun.

h. Menentukan kelas (node daun) yang menunjukkan hasil akhir klasifikasi teks berdasarkan kelas yang dominan.

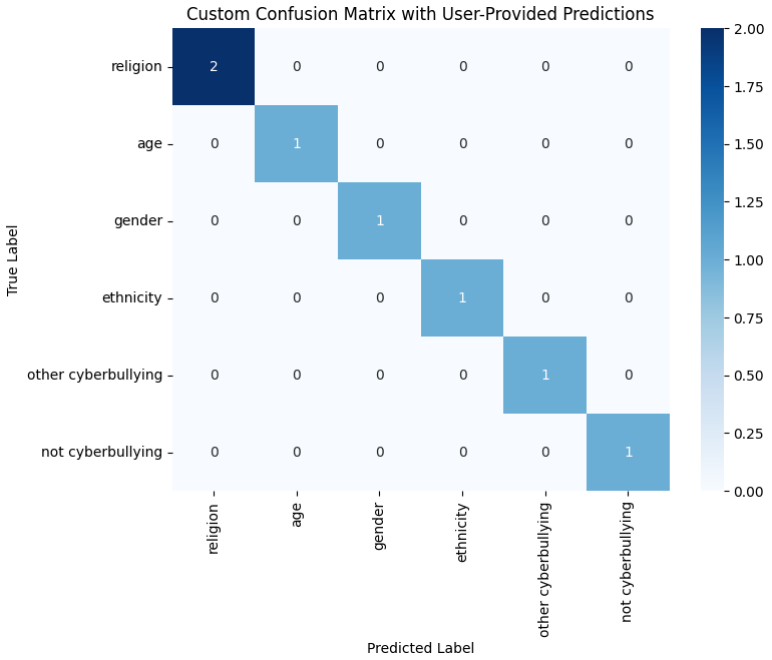
3.6.1 Kombinasi *Grid Search Cross Validation*

Pada Decision Tree, untuk mengoptimalkan performa model dapat dilakukan salah satunya dengan mengukur kedalaman maksimal yang terbaik untuk model, dan teknik yang sering digunakan untuk menentukan parameter terbaik guna meningkatkan performa dari model adalah *Grid Search Cross Validation* (Ayu Firnanda et al., 2025). Teknik ini terbagi menjadi dua bagian dimana *grid search* bertugas mencoba berbagai pilihan kedalaman, dan K-fold Cross Validation bertugas menguji dan memastikan bahwa kedalaman stabil dan akurat bedasarkan data yang diolah oleh model *Decision Tree*.

3.7 Evaluasi

Pada penelitian ini pengujian dilakukan dengan menggunakan *confusion matrix*. *Confusion matrix* memberikan gambaran mengani hasil klasifikasi, sehingga dapat memahami sejauh mana model *Decision tree* berhasil dalam

membedakan antara kasus *cyberbullying*. Dari *confusion matrix*, kita dapat menghitung berbagai metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*, yang memberikan wawasan mendalam tentang performa model. Hasil evaluasi ini dapat digunakan sebagai dasar untuk menyimpulkan bagaimana metode *Decision tree* dengan ekstraksi fitur TF-IDF efektif dalam mengatasi masalah klasifikasi *cyberbullying*. Dengan pemahaman yang lebih baik mengenai kinerja model, penelitian ini dapat memberikan panduan berharga untuk pengembangan sistem klasifikasi *cyberbullying* yang lebih efektif dan efisien di masa depan. Dalam penelitian ini menggunakan evaluasi *confusion matrix* dalam bentuk *multiclass* yang dapat dilihat pada gambar 3.8.



Gambar 3.8 *Confusion Matrix* dari Sampel

Untuk mengukur kinerja model klasifikasi, digunakan beberapa metrik evaluasi yang dirumuskan sebagai berikut:

1. Rumus Accuracy dapat dihitung menggunakan persamaan 3.9

$$\frac{TP + TN}{TP + FP + TN + FN} \quad (3.9)$$

2. Rumus Precission dapat dihitung menggunakan Persamaan 3.10

$$\frac{TP}{TP + FP} \quad (3.10)$$

3. Rumus Recall dapat dihitung menggunakan Persamaan 3.11

$$\frac{TP}{TP + FN} \quad (3.11)$$

4. Rumus f-measure dapat dihitung menggunakan Persamaan 3.12

$$2.\frac{\text{precission. recall}}{\text{precission} + \text{recall}} \quad (3.12)$$

Hasil klasifikasi dapat dianggap benar jika kategori atau tabel yang sebenarnya dari data uji sesuai dengan hasil yang telah diberikan oleh sistem klasifikasi. Dalam penelitian ini, digunakan lebih dari dua kategori atau kelas yang berbeda. Berikut adalah penjelasan istilah yang digunakan dalam *confusion matrix*:

1. *True Positive* (TP) diperoleh dengan jumlah data yang memiliki nilai positif dan diklasifikasikan sebagai data positif.
2. *True Negative* (TN) diperoleh jumlah data yang memiliki nilai *negative* dan diklasifikasikan sebagai data negative.
3. *False Positif* (FP) diperoleh dengan jumlah data yang memiliki nilai *negative* tetapi diklasifikasikan sebagai data positif.
4. *False Negative* (FN) diperoleh dengan jumlah data yang memiliki nilai positif tetapi diklasifikasikan sebagai data *negative*.

3.8 Skenario Pengujian

Pada tahap ini dilakukan skenario pengujian uji coba bagaimana metode *Decision tree* mengolah data uji yang telah disiapkan dengan bobot akhir dari hasil proses metode *Decision tree*, data uji juga didapat dari Kaggle yang sama dengan data latih tetapi data uji ini tidak dimasukkan ke dalam proses pelatihan. Untuk mengetahui kinerja sistem diperlukan perghitungan *accuracy, precision, recall* dan f1-*score*. Recall merupakan perbandingan jumlah dokumen yang relevan yang terambil sesuai dengan *query* yang diberikan dengan total kumpulan dokumen yang relevan dengan *query* (Fahmy Amin, 2022). *Presicion* merupakan perbandingan jumlah dokumen yang relevan terhadap *query* dengan jumlah dokumen yang terambil dari hasil pencarian.

Pembagian dataset akan dilakukan dengan berbagai skenario yang berbeda untuk membandingkan hasil klasifikasi dari masing-masing skenario. Tujuannya adalah untuk menentukan skenario yang menghasilkan hasil klasifikasi terbaik. Pembagian dataset dapat dilihat pada Tabel 3.12, hal ini dilakukan untuk membantu dalam memahami dan mengevaluasi kualitas klasifikasi yang berbeda dalam berbagai konteks atau kondisi yang mungkin terjadi.

Tabel 3.17 Pembagian Dataset

No.	Data Latih	Data Uji
1.	70%	30%
2.	80%	20%
3.	90%	10%

Pengujian kedalaman dari decision tree dilakukan dengan melakukan percobaan berulang hingga memperoleh titik jenuh pada rata rata dan standar deviasi-nya, teknik yang digunakan ini secara umum disebut dengan teknik *grid search* yakni percobaan berulang hingga menemukan kedalaman maksimal yang optimal digunakan oleh model berdasarkan jumlah data yang dilatih.

BAB IV

UJI COBA DAN PEMBAHASAN

4.1 Gambaran Umum Data

Dataset terdiri dari 47.692 data tweet yang sudah dilabeli dengan jenis klasifikasi *cyberbullying* dimana masiing masing label terdiri dari kurang lebih sebanyak 7.948 data tweet, dataset ini disusun diusahakan dengan rata agar data mendekati balance sehingga tidak menimbulkan ketimpangan pada klasifikasi label tertentu dengan tujuan agar pengolahan model terhadap data dapat menghasilkan klasifikasi yang optimal, perincian label dari dataset dapat ditampilkan pada tabel 4.1.

Tabel 4.1 Rincian Dataset Tweet

Label Tweet	Jumlah
Not_Cyberbullying	7945
Gender	7973
Religion	7998
Other_Cyberbullying	7823
Age	7992
Ethnicity	7961

Setelah proses pelabelan dataset dan penentuan jumlah masing masing label tweet, tahap yang selanjutnya dilakukan untuk menyiapkan data adalah tahap *prerprocessing* yakni pengolahan teks melalui normalisasi, *tokenizing*, *stemming*, dan *stopword removal* sehingga menghasilkan data yang mudah diolah oleh model.

4.2 Hasil Preprocessing dan Ekstraksi Fitur

Melalui tahap *preprocessing* yang terdiri dari tahap *stopword removal*, tahap pembersihan *emoji* yang muncul pada data *text tweet*, menghapus kata kata yang menggunakan singkatan yang umum digunakan pada teks *tweet*, menghapussimbol *hashtag* dan simbol-simbol lain pada *tweet*, menghilangkan teks spasi berlebih, setelah semua teks dinyatakan bersih dari simbol, karakter, dan *emoji* dapat dilanjutkan ke tahap selanjutnya. Tahap *tokenizing* yakni tahap pemecahan teks kalimat kedalam bentuk per kata, tidak selain itu juga perlu dilakukannya proses *stemming* untuk menghubah kata berimbuhan menjadi kata dasar, serupa dengan proses *stemming*, *lemmatization* membentuk proses mengubah teks kalimat yang lebih kompleks menjadi kata dasar yang mudah diketahui makna dan artinya sehingga memudahkan proses untuk menentukan apakah teks tweet termasuk pada *cyberbullying* maupun *non-cyberbullying*, adapun cuplikan hasil *preprocessing* tweet yang diolah menjadi teks sederhana dapat ditampilkan pada tabel 4.2.

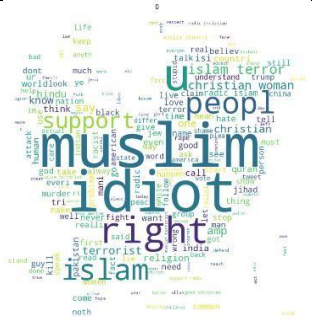
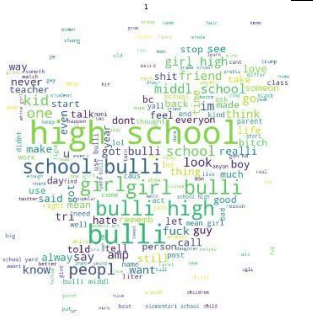
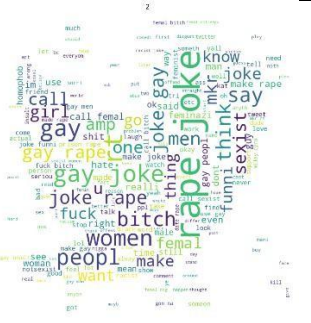
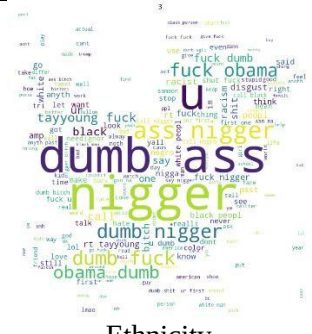
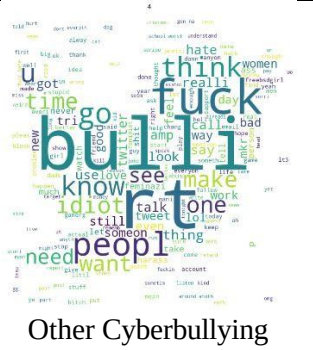
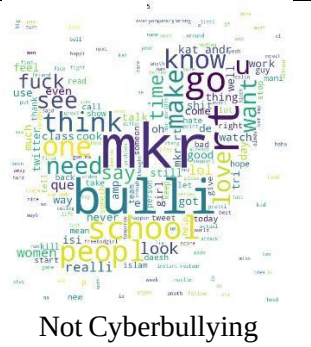
Tabel 4.2 Cuplikan Hasil Preprocessing Teks Tweet

Teks Asal	Teks Hasil *Preprocessing*
In other words #katandandre, your food was crapilicious! #mkr	word katandandr food crapilici mkr
Why is #aussietv so white? #MKR #theblock #ImACelebrityAU #today #sunrise #studio10 #Neighbours #WonderlandTen #etc	aussietv white mkr theblock today sunris studio10 neighbour wonderlandten etc

@XochitlSuckkks a classy whore? Or more red velvet cupcakes?	classi whore red velvet cupcak
@Jason_Gio meh. :P thanks for the heads up, but not too concerned about another angry dude on twitter.	meh p thank head concern anoth angri dude twitter
@RudhoeEnglish This is an ISIS account pretending to be a Kurdish account. Like Islam, it is all lies.	isi account pretend kurdish account like islam lie

Hasil teks dari tabel 4.2 menunjukkan bahwa teks *tweet* berhasil diolah menjadi kata sederhana. Tahap berikutnya yang dilakukan adalah penyaringan *tweet* dimana hal ini bertujuan untuk menhindari *tweet* spam yang hanya berisi kurang dari 3 kata dalam satu *tweet*, dan menghindari anomali seperti teks *tweet* dengan jumlah kata lebih dari 100 kata dalam satu *tweet*, penyaringan ini juga bertujuan meminimalkan teks statistik TF-IDF menguji teks *tweet* secara berlebihan, sehingga dilakukan penyaringan jumlah kata dalam teks *tweet* dengan batasan lebih dari 3 kata dan kurang dari 100 kata. Kemudian data teks yang didapatkan masing masing kelasnya akan di petakan pada wordcloud untuk menunjukkan kata yang sering muncul dalam setiap kelasnya.

Tabel 4.3 Visualisasi *WordCloud* Setiap Kelas

 Religion 7.893 Data	 Age 7.830 Data	 Gender 7.311 Data
 Ethnicity 7.707 Data	 Other Cyberbullying 6.371 Data	 Not Cyberbullying 4.669 Data

4.2.1 Pengolahan TF-IDF

Pada tahap TF-IDF data teks *tweet* mengalami pembobotan berdasarkan frekuensi kemunculan kata dalam sebuah teks, berdasarkan pengolahan TF-IDF ini pada sistem klasifikasi teks *tweet cyberbullying* diperoleh ukuran matriks dengan (41782, 34806) yang berarti dari sebanyak 41782 teks *tweet* yang diolah, diperoleh sebanyak 34806 kata unik yang digunakan dalam teks tweet. Untuk mengukur nilai statistic TF-IDF diperlukan untuk mengetahui nilai TF (*Term Frequency*) dan pada bagian ini digunakan fungsi CountVectorizer() dan atribut *fit_transform* pada data teks yang sudah disederhanakan melalui *preprocessing*, fungsi ini nantinya menyimpan frekuensi kemunculan kata dalam setiap teks *tweet* kedalam *bag of*

words, pada tabel 4.4 menunjukkan nilai dari *Term Frequency* untuk beberapa sampel kata.

Tabel 4.4 Cuplikan nilai TF terhadap kata dalam teks *tweet*

Kata dengan nilai TF rendah		Kata dengan nilai TF tinggi	
Kata	Nilai TF	Kata	Nilai TF
Bulli	10695	zoo	1
school	8977	zionist	1
like	6139	zay	1
dumb	5124	yup	1

Selanjutnya adalah mengetahui nilai IDF (*Inverse Document Frequency*) dimana pada nilai ini dihitung frekuensi kelangkaan kata yang muncul pada teks, sehingga jika ada kata yang hanya muncul pada teks *tweet* tertentu makan nilai bobotnya besar. Berdasarkan dataset teks *tweet* yang telah diolah didapatkan cuplikan nilai IDF pada kata tertentu terdapat pada tabel 4.5.

Tabel 4.5 Cuplikan nilai IDF terhadap kata dalam teks *tweet*

Kata dengan nilai IDF rendah		Kata dengan nilai IDF tinggi	
Kata	Nilai IDF	Kata	Nilai IDF
bulli	2.496151	ztraight	10.947098
school	2.663098	ztupidd	10.947098
like	3.042394	zucchini	10.947098
dumb	3.152480	zuckerberg	10.947098

Dari tabel 4.3 ditunjukkan beberapa kata yang memiliki nilai IDF kecil yang berarti umum muncul pada teks *tweet* yang diolah, dan juga terdapat beberapa kata yang memiliki nilai IDF tinggi yang berarti kata tersebut unik atau jarang muncul pada data teks *tweet* yang diolah. Setelah mengetahui nilai TF dan IDF, selanjutnya adalah menghitung nilai TF-IDF dengan cara dikalikan dan menghasilkan nilai angka yang nantinya bisa diolah dengan *tree decision* untuk mengetahui keputusan apakah termasuk klasifikasi *cyberbullying* atau *non-cyberbullying* berikut adalah

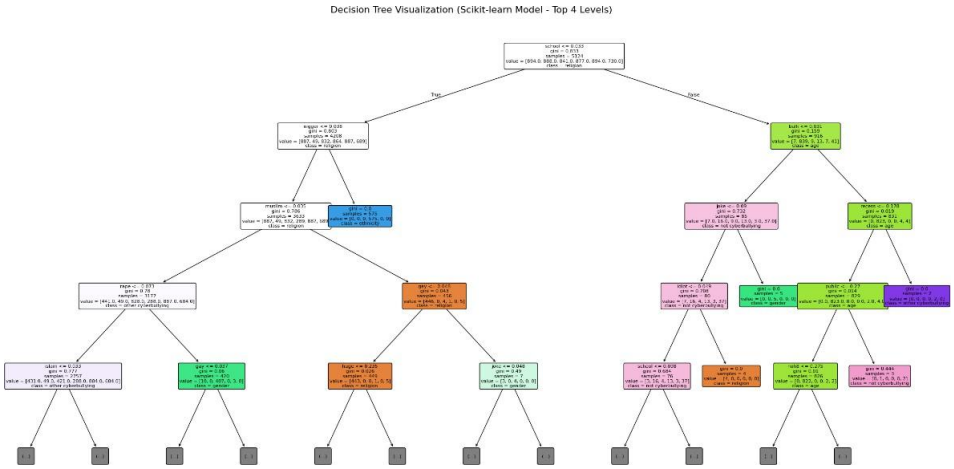
cuplikan tabel untuk menapilkan implementasi perhitungan TF-IDF terhadap dataset.

Tabel 4.6 Cuplikan Sampel Perhitungan TF-IDF

Kata	Total *Term Frequency*	Inverse Document Frequency	Nilai TF-IDF
nigga	600	5.348676	0.838407
bulli	10695	2.496151	0.770302
rape	4323	3.371257	0.741877
kid	1202	4.691348	0.750387
fuck	7533	2.883877	0.792981
muslim	4723	3.357509	0.567164
gay	4181	3.377944	0.628499
christian	2221	4.085386	0.589200

4.2.2 Decision Tree

Pada tahap ini dilakukan implementasi komputasi decision tree terhadap masing-masing kata yang telah melalui proses statistik TF-IDF, berikutnya adalah proses algoritma *decision tree* untuk memberikan keputusan pada setiap kata untuk mengklasifikasifikasikannya kedalam sentimen label, berikut ini merupakan contoh struktur pohon yang ditampilkan oleh model decision tree dengan menggunakan fitur kata "*school*"



Gambar 4.1 Diagram Pohon *Decision Tree*

Gambar 4.1 merupakan gambar diagram pohon dari *decision tree* dengan, dimana pada visual ini dilakukan pengujian sample terhadap kata *school* untuk menentukan kelas sentimennya, pada decision tree yang ditunjukkan pada gambar 4.1, diketahui bahawa *root node* (node akar) adalah kata school, dengan keputusan apabila nilai TF-IDF dari fitur kurang sama dengan *threshold* (True) maka akan turun ke *node leaf* kiri dan apabila lebih dari nilai *threshold* (False) maka akan turun ke *node leaf* kanan satu kali pada diagram diatas ditunjukkan bahwa balon biru merupakan *node* internal dimana terjadi perbandingan keputusan antara nilai TF-IDF dengan *threshold*, dan pada *node leaf* sebelah kanan merupakan hasil prediksi sentiment dari masing masing node internal sebelumnya, berdasarkan gambar 4.1 jika diperhatikan pada sisi kanan terdapat dua output yang menunjukkan bahwa kata *school* diprediksikan ke dalam sentimen *not cyberbullying* dan *other cyberbullying*. Berikut ini merupakan skema aturan *decision tree* dengan memanfaatkan TF-IDF berdasarkan batas kedalaman maksimal yang ditunjukkan pada Tabel 4.7, dimana pada tabel ini menunjukkan tahapan penentuan keputusan sebuah kata berdasarkan nilai TF-IDF dengan kedalaman yang dibutuhkan, pada tabel ini sample kata yang digunakan adalah "*Rebecca Black Drops Out of School Due to Bullying*" dengan label not *cyberbullying*, kemudian teks di bersihkan menjadi "*rebecca black drop school due bulli*" lalu diproses berdasarkan kedalaman dan perbandingan TF-IDF dengan nilai threshold nya untuk memperoleh keputusan prediksi.

Tabel 4.7 Tahapan Keputusan bedasarakan Kedalaman dan TF-IDF

Kedalaman	Tipe Node	Fitur Kata	Threshold	Nilai TF-IDF	Keputusan
0	Internal	school	0.0413	0.1943	Is 'school' <= 0.0413 (TF-IDF

					value)? -> No (Go Right)
1	Internal	bulli	0.0200	0.1821	Is 'bulli' <= 0.0200 (TF-IDF value)? -> No (Go Right
2	Internal	rebecca	0.5831	0.6062	Is 'rebecca' <= 0.5831 (TF-IDF value)? -> No (Go Right)
3	Leaf	Final Prediction	-	-	Predict: not cyberbullying

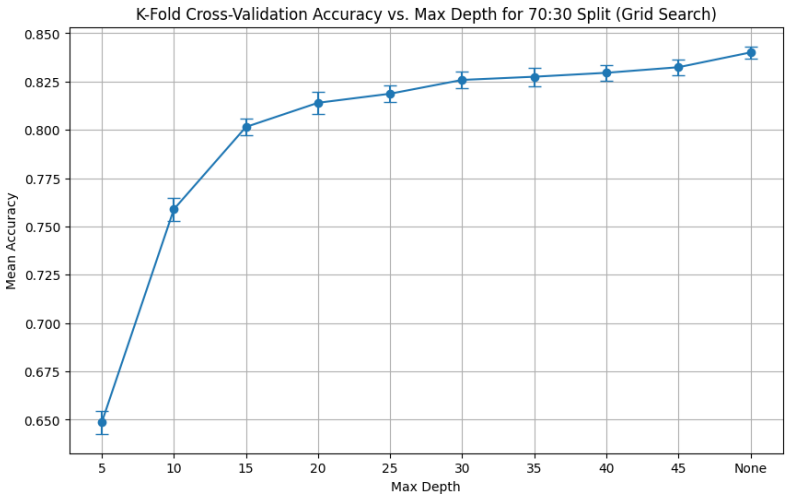
4.3 Peforma Model

Pada penelitian ini model *decision tree* yang digunakan memiliki model sturktur pohon dengan batas kedalaman yang diukur dengan teknik *grid search*, untuk menguji tiga skenario pembagian data uji dan data latih sesuai pada tabel 3.13, dimana pada tahap ini evaluasi dilakukan untuk menguji keberhasilan model terhadap skenario yang ada dengan parameter pengujian recall, presisi, skor f1, dan akurasi, pengujian masing-masing skenario dilakukan dengan pembagian data terlebih dahulu untuk mengetahui jumlah data yang akan dilatih, dan untuk mengantisipasi jumlah data yang tidak stabil digunakan teknik SMOTE (*Synthetic Minority Over-sampling Technique*) yakni teknik untuk memanipulasi jumlah data (*oversampling*) hal ini dilakukan sebab setelah melakukan pembersihan teks data twit diketahui bahwa total data tidak seutuhnya 47692 teks melainkan hanya berjumlah 41782, maka dari itu penting dilakukan teknik *oversampling* ini, tahapan berikutnya akan dilakukan uji coba dengan grid search dan juga kombinasi dengan *k-fold cross validation* pada masing-masing skenario untuk mengetahui batas maksimal ke dalam masing-masing skenario, dan selanjutnya dilakukan analisis

dengan *confusion matrix* dan perbandingan hasil rata rata keempat metrik untuk masing-masing skenario.

4.3.1 Skenario 1

Pada skenario pertama ini dilakukan pembagian data latih sebanyak 70% (±4800 data teks tweet) dan data uji (± 2000 data teks tweet), diketahui berdasarkan pembagian data ini melalui teknik *grid search* dan K-*fold cross validation* diperoleh kedalaman terbaik dengan akurasi 0.8401, pada gambar 4.2 ditunjukkan grafik garis yang menunjukkan nilai rata-rata akurasi terhadap kedalaman maksimum.



Gambar 4. 2 Grafik Max Depth Skenario 70:30

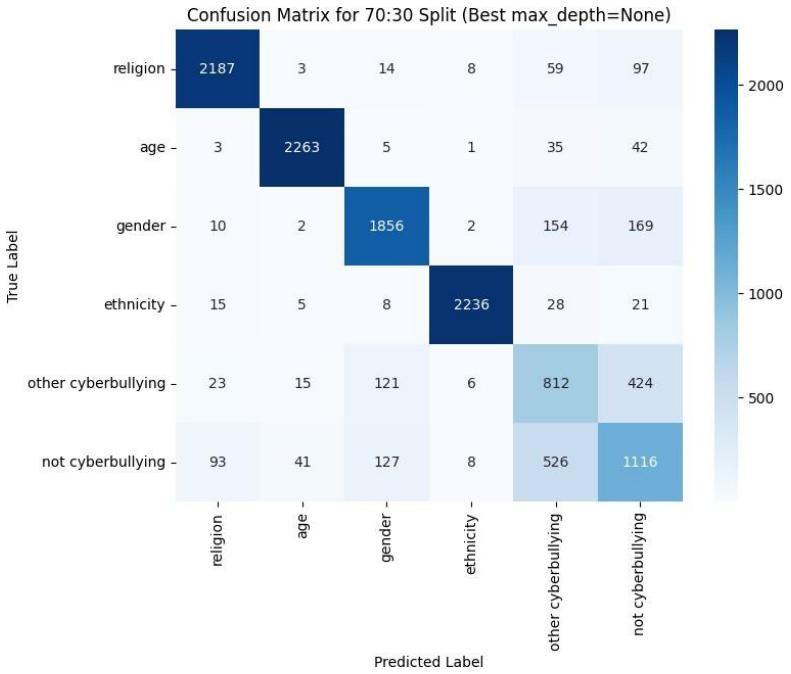
Berdasarkan gambar 4.2, ditunjukkan bahwa nilai rata-rata akurasi berada pada puncak grafik di mana pada hasil nilai diperoleh rata-rata akurasi sebesar 0.8401, kemudian dari pembagian ini akan diukur bagaimana tingkat performa dari model *decision tree* terhadap skenario perbandingan 70:30.

Tabel 4.8 Performa Seknario 70:30 terhadap sentimen

Sentimen	Recall	Presisi	Skor f1
Religion	0.92	0.94	0.93

Age	0.96	0.97	0.97
Gender	0.85	0.87	0.86
Ethnicity	0.97	0.99	0.98
Other Cyberbullying	0.58	0.50	0.54
Not Cyberbullying	0.58	0.60	0.59

Tabel 4.8 menunjukkan nilai performa model terhadap skenario 1 dimana dari tabel ini diperoleh untuk label *religion*, *age, gender*, dan *ethnicity* memilik nilai preisisi, recall, dan skor f1 yang cukup tinggi di mana berada diantara 80% hingga 90% sedangkan untuk label *other cyberbullying* dan *not cyberbullying* memiliki nilai metrik yang cukup rendah yakni berada dantara 50% hingga 60% dan dari tabel dapat ditarik informasi bahwa pada perbandingan skenario ini model kurang mampu mendeteksi sentimen *not bullying* dan *other cyberbullying* hal ini didasari dari nilai rata-rata metrik yang cukup rendah, sebaliknya model mampu mendeteksi sentimen cyber*bullying* lainnya secara baik untuk masing masing labelnya, berikut ditampilkan pada gambar 4.3 untuk memperjelas jumlah data prediksi masing masing sentimen pada skenario perbandingan 70:30 dengan menggunakan pemetaan *confusion matrix*.

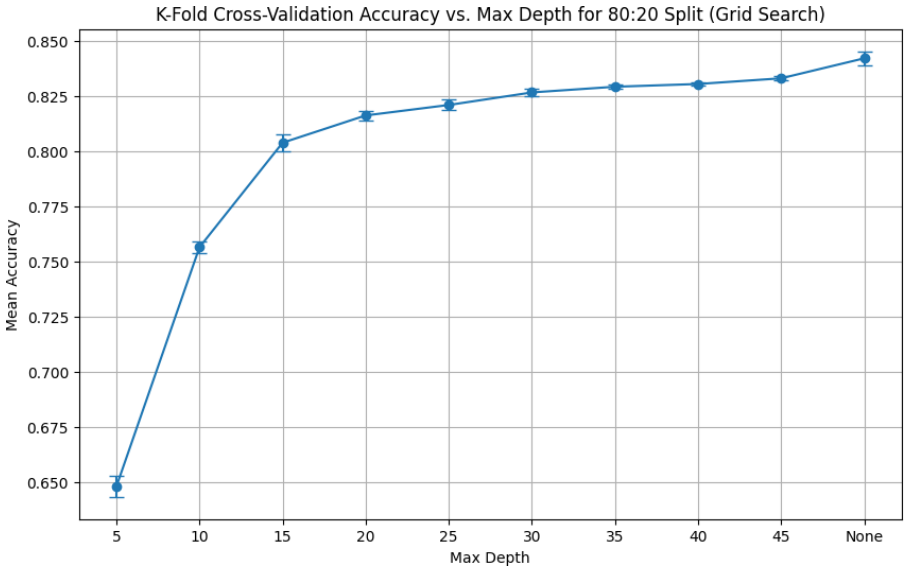


Gambar 4.3 *Confusion Matrix* Skenario 1 70:30

Gambar 4.3 menunjukkan pemetaan confusion matrix untuk skenario perbandingan data tweet 70:30 dari ±7000 data tweet, dan dari gambar ini diperoleh informasi bahwa terdapat beberapa kesalahan prediksi teks dimana beberapa hasil prediksi teks berbeda dengan teks aslinya, dari gambar 4.3 dapat diketahui bahwa prediksi sentimen label *other cyberbullying* dan *not cyberbullying* mengalami kesalahan prediksi di mana *other cyberbullying* sebanyak 424 data dan *not cyberbullying* sebanyak 526 data.

4.3.2 Skenario 2

Skenario kedua ini dilakukan pembagian data latih dan data uji sebesar 80:20 yakni kurang lebih sebanyak ± 5570 data latih dan ± 1392 data uji pada tiap kelas sentimen, melalui teknik *cross validation* dan *grid search* dipeeroleh kedalaman maksimum yang dapat ditunjukkan pada gambar 4.4.



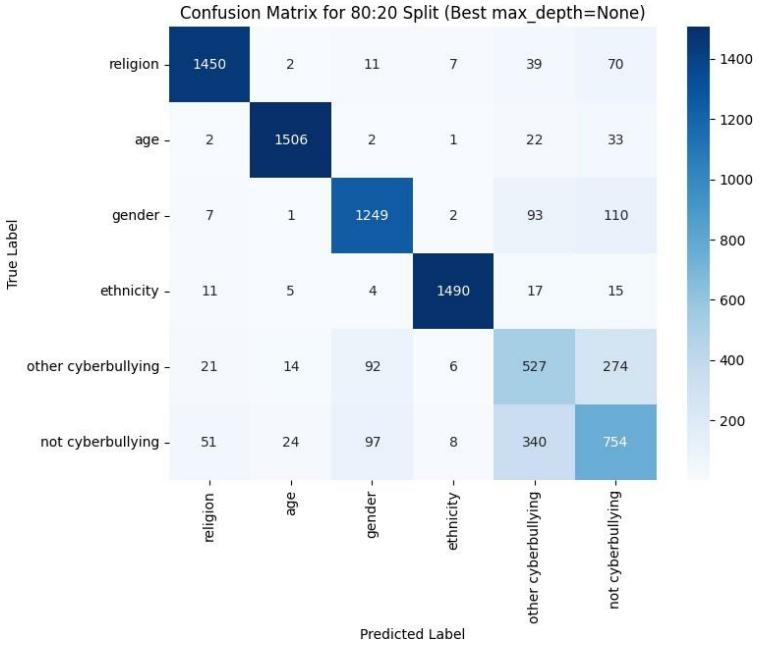
Gambar 4.4 Grafik Max Depth Skenario 80:20

Berdasarkan gambar 4.4 diperoleh kedalaman maksimum untuk model dengan skenario perbandingan data ini memperoleh akurasi rata-rata sebesar 0,816, selanjutnya dilakukan implementasi *decision tree* pada skenario ini dan berikut performa yang dihasilkan dari perbandingan data terhadap model ditunjukkan pada tabel 4.9.

Tabel 4.9 Performa Skenario 80:20 tedahap sentimen

Sentimen	Recall	Presisi	Skor f1
Religion	0.92	0.94	0.93
Age	0.96	0.97	0.97
Gender	0.85	0.86	0.86
Ethnicity	0.97	0.98	0.98
Other Cyberbullying	0.56	0.51	0.53
Not Cyberbullying	0.59	0.60	0.60

Tabel 4.9 menunjukkan hasil performa model terhadap skenario 2 dimana pada tabel ini diperoleh informasi bahwa sentimen *other cyberbullying* dan *not cyberbullying* juga memiliki terendah tepatnya pada metrik presisi dan skor f1 dengan nilai sama dengan kurang dari 60%, dan sebaliknya untuk label *cyberbullying* yang lainnya masih memiliki nilai matrik yang lebih tinggi. Berikut juga akan ditampilkan gambar pemetaan *confusion matrix* dari perbandingan skenario 80:20 untuk memperjelas bagaimana model mampu memprediksi teks tweet terhadap kelas sentimen.



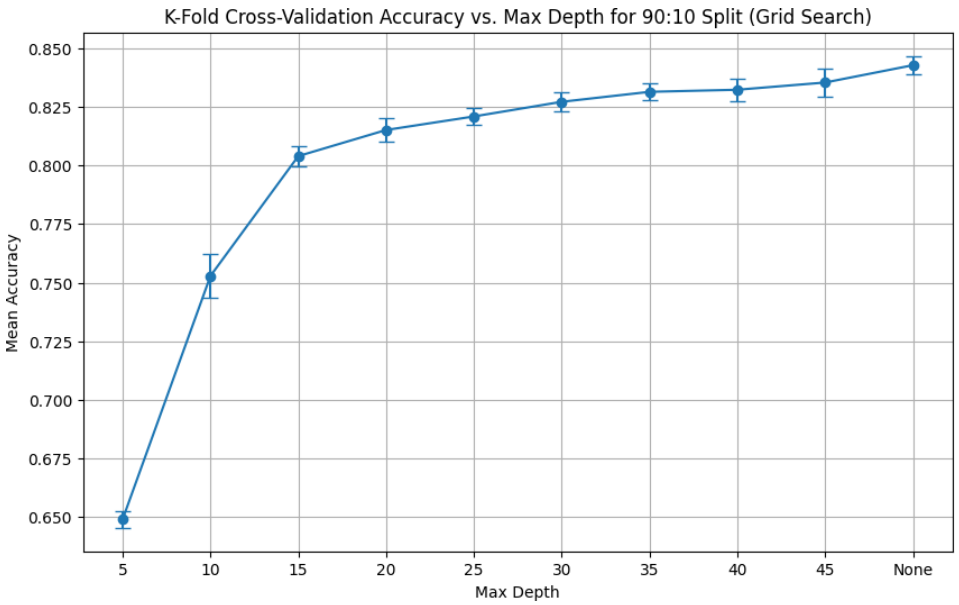
Gambar 4.5 *Confusion Matrix* Skenario 2 80:20

Gambar 4.5 menunjukkan pemetaan *confusion matrix* untuk skrenario kedua yakni skenario pembagian data latih dan data uji sebesar 80:20, dimana dari hasi pemetaan diperoleh dari hasil pemetaannya memiliki prediksi kesalahan yang lebih kecil dibandingkan dengan skenario perbandingan 70:30 namun dari perbedaan nilai kesalahan prediksi ini tidak sepenuhnya sesuai sebab perbedaan

skenario membuat perbedaan jumlah data uji yang ditampilkan pada peta confusion matrix, dari gambar 4.5 dapat dilihat bahwa prediksi *other cyberbullying* dan *not bullying* masih mengalami kekeliruan dimana kesalahan prediksi terdapat pada label gender yang diprediksi sebagai label *other cyberbullying* sebanyak 92 data dan *not cyberbullying* sebanyak 97 data, kemudian masing-masing pada prediksi other cyber bullying sebanyak 340 data dan not cyber bullying sebanyak 247 data..

4.3.3 Skenario 3

Pada skenario ketiga ini dilakukan pembagian data latih dan data uji sebanyak kurang lebih 6267 data teks tweet sebagai data latih pada tiap sentimen dan kurang lebih 696 data teks tweet pada tiap sentimen sebagai data uji, selanjutnya melalui teknik *cross validation* dan *grid search* diperoleh kedalaman maksimum yang ditunjukkan pada gambar grafik 4.6.



Gambar 4.6 Grafik Max Depth Skenario 90:10

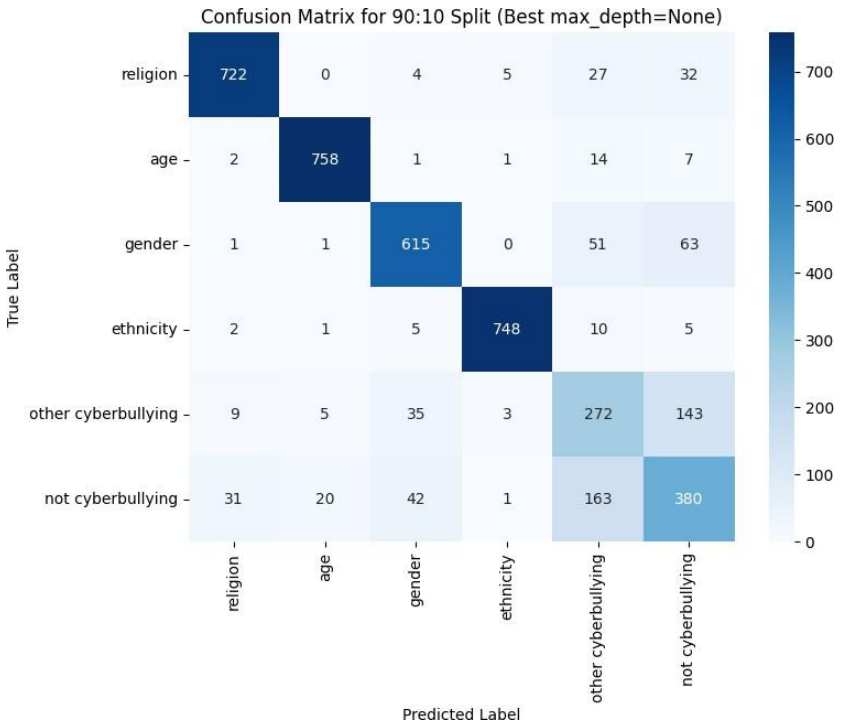
Berdasarkan gambar 4.6 ditunjukkan pada grafik dimana kedalaman maksimum untuk skenario 3 dengan nilai rata-rata akurasi sebesar 0.8427, dan

setelah diimplementasikan pada model *Decision Tree* berikut performa yang dihasilkan dari perbandingan 90:10 terhadap model dapat ditunjukkan pada tabel 4.10.

Tabel 4.10 Performa Skenario 90:10 terhadap setimen

Sentimen	Recall	Presisi	Skor f1
Religion	0.91	0.94	0.93
Age	0.97	0.97	0.97
Gender	0.84	0.88	0.86
Ethnicity	0.97	0.99	0.98
Other Cyberbullying	0.58	0.51	0.54
Not Cyberbullying	0.60	0.60	0.60

Tabel 4.10 menunjukkan hasil performa model terhadap skenario ketiga yakni 90:10 di mana pada tabel ini ditunjukkan informasi bahwa terdapat label sentimen yang memiliki skor terbaik yakni pada sentimen dengan label *age* dengan rata-rata di atas 90%, dan sebaliknya pada performa skenario ini juga masih menunjukkan nilai metrik yang cukup rendah pada label sentimen o*ther cyberbullying* tepatnya pada metrik presisi dan skor f1, kemudian pada label sentimen *not cyberbullying* yang menunjukkan nilai metrik sentimen terendah jika dibandingkan dengan label sentimen yang lain. Untuk mengetahui kesalahan prediksi yang terdapat pada skenario 3 ini dengan perbandingan data 90:10, dapat ditunjukkan pada pemetaan *confusion metrix* yang ditunjukkan pada gambar 4.7.



Gambar 4.7 *Confusion Matrix* Skenario 3 90:10

Confusion matrix dengan perbandingan skenario data *tweet* sebesar 90:10 menunjukkan bahwa bias model terhadap data semakin kecil, dimana dari pemetaan yang ditampilkan pada gambar 4.7 dapat diketahui bahwa kesalahan-kesalahan prediksi pada sentimen tidak terlalu besar jumlah kesalahannya jika dibandingkan dengan skenario 1 dan skenario 2, jumlah data yang mengalami kesalahan prediksi pada perbandingan 90:10 dengan kedalaman maksimum yang telah ditentukan dengan teknik grid search dan k-fold corss validation, ditunjukkan pada label sentimen *other cyberbullying* dengan 163 data yang seharusnya termasuk klasifikasi teks *not cyberbullying* dan *not cyberbullying* sebanyak 143 data yang seharusnya berlabel *other cyberbullying*.

4.3.4 Akurasi Model

Pada tahapan ini akan diambil rata-rata akurasi dari setiap skenario terhadap model *decision tree* yang digunakan dengan kedalaman maksimum yang sudah diukur dengan kombinasi *grid search* dan *cross validation* untuk masing-masing skenario yang berbeda. Berikut ini disajikan tabel yang akan menunjukkan nilai akurasi dari masing-masing skenario berdasarkan kedalaman maksimum dari masing-masing skenario.

Tabel 4.11 Hasil akurasi masing-masing skenario

Skenario Pengujian Data Tweet	Nilai Akurasi
70% Data Latih 30% Data Uji	0.8353
80% Data Latih 20% Data Uji	0.8347
90% Data Latih 10% Data Uji	0.8363

Dari tabel 4.11 ditunjukkan hasil akurasi dari masing-masing skenario pengujian data *tweet*, dan dari nilai akurasinya dapat ditunjukkan perbandingan 90:10 memiliki nilai akurasi tertinggi sebesar 83,63%% sedangkan untuk perbandingan 70:30 dan 80:20 hanya memperoleh nilai akurasi sebesar 83,53% dan 83,47% saja.

4.4 Pembahasan

Pengujian model *decision tree* dengan mengkombinasikan *cross validation* dan *grid search* untuk mengetahui kedalaman masimal dari model berdasarakan kesesuain perbandingan data pada tiap skenario menghasilkan akurasi terbaik pada skenario 3, yakni menggunakan perbandingan 90:10 dan kedalaman terbaik, menghasilkan akurasi yang cukup tinggi sebesar 0,8363atau 83,63%%. Hasil nilai

akurasi ini membuktikan bahwa model cukup mampu dalam mengklasifikasikan teks data *tweet* dengan enam kelas label sentimen meskipun terdapat kesalahan prediksi dalam mengklasifikasikan label *other cyberbullying* dan *not cyberbullying*.

Pada kelas *other cyberbullying* dan *not cyberbullying* menunjukkan nilai rata-rata akurasi yang terendah, hal ini disebabkan oleh ekstraksi fitur kata pada kelas *other cyberbullying* dan not *cyberbullying* memiliki kesamaan kata yang sering muncul dalam kedua kelas sentimen tersebut, hal ini ditunjukkan pada tabel 4.3 dimana pada visualisasi seringnya kata muncul pada tiap kelas, untuk *other cyberbullying* dan *not cyberbullying* memiliki kesamaan kata contohnya "bulli", "like", "get", "go", dan "would", selain itu juga jumlah kata setelah *preprocessing* menunjukkan untuk kedua kelas ini memiliki jumlah data yang rendah di mana untuk label sentimen *other cyberbullying* memiliki jumlah data terendah yakni hanya sebesar 4.669. Berikut ini disajikan Tabel 4.12 yang menunjukkan jumlah kata yang sering muncul pada setiap sentiment dan juga jumlah data pada setiap sentimen guna menunjukkan jumlah data khususnya pada sentimen other cyberbullying dan not cyberbullying.

Tabel 4.12 Jumlah kata yang sering muncul dan jumlah data

Sentimen	Kata yang Sering Muncul	Jumlah Data Setelah Preprocessing
religion	muslim (4603), idiot (3066), islam (2429), christian (2126), terrorist (1373), right (1287), like (1274), support (1247), terror (1179), radic (1101)	7893

age	bulli (8861), school (8386), high (4855), girl (4613), like (2075), get (1051), one (1043), peopl (933), got (840), kid (808)	7830
gender	joke (5146), rape (4045), gay (3831), call (1393), make (1278), rt (1184), bitch (1122), femal (1102), peopl (985), like (970)	7311
Ethnicity	fuck (5831), nigger (5396), dumb (4950), ass (2210), black (2123), white (1536), call (1375), peopl (1177), rt (1156), obama (1086)	7708
other cyberbullying	bulli (723), rt (659), fuck (446), like (389), get (372), peopl (337), go (269), think (223), would (221), idiot (220)	4669
not cyberbullying	mkr (1465), bulli (1023), rt (748), get (410), like (401), go (368), school (338), would (276), kat (276), peopl (263)	6371

Kedalaman maksimum yang digunakan pada penelitian klasifikasi teks cyberbullying dengan menggunakan metode *Decision Tree* ini sangat berpengaruh secara signifikan, sebab berdasarkan data skenario yang diperoleh pada grafik garis kedalaman maksimum, apabila nilai kedalamannya terlalu rendah atau dangkal maka rata-rata akurasi yang dihasilkan juga kurang maksimal dan juga beresiko terjadinya underfitting yakni performa data latih dan uju sama sama rendah sebab data yang digunakan cukup banyak yakni 47.692 sebelum tahap preprocessing dan 41.781 setelah tahap preprocessing, sedangkan apabila kedalaman yang digunakan terlalu tinggi atau bebas maka akan berdampak pada lamanya model membuat cabang hingga tingkat kata tertentu pada teks tweet hal ini juga berisiko terjadinya overfitting model terhadap data dimana performa data latih lebih baik dari pada data

uji secara signifikan, maka pada penelitian ini dengan melihat jumlah data yang besar dan pembagian data per kelas yang kurang seimbang penting dilakukan untuk menentukan batas kedalaman maksimum yang optimal pada setiap skenario pembagian data yang ada. Pada Gambar 4.2, Gambar 4.4, dan Gambar 4.6, terlihat bahwaa nilai rata-rata akurasi mengalami peningkatan yang cukup signifikan pada tahap awal saat nilai kedalaman maksimum dinaikkan, yang menunjukkan bahwa model mulai mampu mempelajari pola kompleks dari data. Namun, setelah mencapai titik kedalaman maksimum yang optimal, peningkatan akurasi mulai melambat dan cenderung stabil, atau bahkan dalam beberapa kasus bisa mengalami penurunan sedikit akibat *overfitting*—kondisi di mana model menjadi terlalu "hafal" dengan data latih namun kehilangan kemampuan generalisasi pada data uji. Oleh karena itu, penggunaan *max depth* yang terlalu tinggi tidak selalu menjamin akurasi yang lebih baik; pemilihan nilai *max depth* yang tepat sangat krusial agar model dapat mencapai keseimbangan antara kemampuan memprediksi dan kompleksitas model sehingga performa klasifikasi tetap konsisten.

Rata-rata akurasi 83,63% yang diperoleh pada skenario 90:10 menunjukkan bahwa model *decision tree* mampu dalam memprediksi data teks *tweet* dengan sentimen label “*age*”, “*gender*”, “*ethnicity*”, dan “*religion*”, sedangkan untuk memprediksi sentimen “*other cyberbullying*” dan “*not cyberbullying*” masih kurang mampu dalam mengklasifikasikan data teks *tweet.* Hasil yang didapatkan ini sejalan dengan apa penelitian terkait yang menggunakan metode *Decision Tree* di mana metode ini mampu untuk memprediksi data dengan jumlah yang besar dan banyak kelas dengan cukup tinggi. Perbedaan rata-rata akurasi yang lumayan besar

jika dibandingkan dengan penelitian yang dilakukan oleh Zaenal (2021) disebabkan oleh keberagaman data teks yang berbeda, di mana pada penelitian ini menggunakan jumlah data yang lebih banyak yakni sebanyak 47.692 data dengan 6 kelas dengan pembagian tiap kelas memiliki data sebanyak 7.948 data teks *tweet*. Meskipun demikian setelah melalu *preprocessing* data teks tweet mengalami pengurangan jumlah data di mana pada kelas *other cyberbullying* memiliki jumlah yang paling kecil yakni sebanyak 4.669 data teks tweet saja, disusul dengan sentimen not cyberbullying sebanyak 6.371 data, sentimen gender sebanyak 7.311 data, sentimen ecthicity sebanyak 7.708 data, sentimen age sebanyak 7.830 data, dan sentimen religion sebanyak 7.893 data. Perbedaan jarak antara sentimen *other cyberbullying* dan *not cyberbullying* dengan sentimen lainnya yang cukup jauh ini tentunya berdampak kepada model saat melakukan klasifikasi teks tweet berdasarkan masing-masing skenario pula, berikut ini disajikan tabel 4.13 untuk mengetahui empat metrik penilaian untuk masing masing skenario.

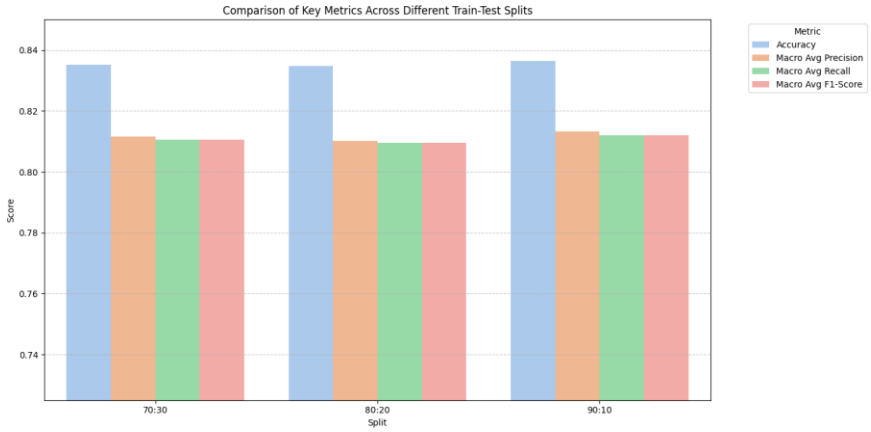
Tabel 4.13 Perbandingan Skenario

Sentimen	Skenario 1 70:30				Skenario 2 80:20				Skenario 3 90:10			
	Recall	Presisi	F1	Acc	Recall	Presisi	F1	Acc	Recall	Presisi	F1	Acc
Religion	0.92	0.94	0.93	0.835	0.92	0.94	0.93	0.83	0.91	0.94	0.93	0.83
Age	0.96	0.97	0.97		0.96	0.97	0.97		0.97	0.97	0.97	

Gender	0.85	0.87	0.86		0.85	0.86	0.86		0.84	0.88	0.86	
Ethnicity	0.97	0.99	0.98		0.97	0.98	0.98		0.97	0.99	0.98	
Other Bullying	0.58	0.50	0.54		0.56	0.51	0.53		0.58	0.51	0.54	
Not Bullying	0.58	0.60	0.59		0.59	0.60	0.60		0.60	0.60	0.60	
Rata Rata	0.810	0.811	0.81	0.835	0.81	0.81	0.80	0.83	0.81	0.813	0.812	0.83

Dari hasil tabel perbandingan skenario terhadap rata-rata metrik yang ditunjukkan pada tabel 4.13, diperoleh bahwa nilai untuk masing-masing metrik pada setiap skenario memiliki perbedaan angka yang tidak terlalu jauh, sehingga hal ini agak sulit untuk diambil kesimpulan mana skenario yang memiliki nilai rata-rata metrik yang terbaik. Berdasarkan tabel 4.13 dapat dilihat pada nilai masing-masing rata-rata metrik memiliki desimal kepala tujuh hingga delapan, dan jika diamati nilai-rata metrik pada skenario tiga yakni pada perbandingan 90:10 memiliki nilai yang cukup tinggi jika dibandingkan dengan skenario lainnya, rata-rata skor f1 yang diperoleh pada skenario 3 sebesar 0,812, rata-rata *recall* sebesar 0,812, rata-rata presisi sebesar 0,813, dan rata-rata akurasi sebesar 83,63%. Berdasarkan nilai rata-rata *recall* juga dapat diketahui bahwa model decision tree mampu untuk mendeteksi semua teks *tweet cyberbullying* yang ada pada data sebenarnya, berdasarkan nilai *recall* sebesar 81,2% menunjukkan model cukup mampu dalam mendeteksi teks *bullying*, meskipun masih terdapat 18,8% teks *bullying* yang tidak terdeteksi, nilai metrik presisi menunjukkan bahwa ketepatan model dalam mengidentifikasi teks cyberbullying sebesar 81,3%, sehingga berdasarkan gabungan dari presisi dan recall menunjukkan rata-rata skor f1 sebesar

18,7 % selain itu untuk mengantisipasi terjadinya overfitting peneliti melakukan perbandingan antara rata-rata akurasi dari hasil data train dan hasil rata-rata akurasi dari data latih, dimana hasil data rata-rata akurasi sebesar 84,27%, sendangkan hasil rata-rata akurasi dari data uji dengan perbandingan 90:10 sebesar 83,63%, hal ini menunjukkan perbedaanya sebesar satu persen yang berarti tidak menunjukkan terjadinya overfitting pada model decision tree yang diimplementasikan pada skenario 90:10. Berikut ini ditampilkan gambar visualisasi untuk memperlihatkan dengan jelas perbedaan nilai dari masing-masing rata-rata metrik pada setiap skenario dengan masing-masing kedalam maksimalnya.



Gambar 4.8 Performa rata-rata metrik pada skenario

Gambar 4.8 menunjukkan hasil pengujian performa dari metode *Decision Tree* dengan ekstraksi fitur menggunakan TF-IDF dalam mengklasifikasikan teks *cyberbullying* pada dataset teks tweet menunjukkan nilai yang signifikan dalam mendeteksi teks dengan konten negatif. Keberhasilan sistem klasifikasi ini sejalan dengan firman Allah SWT, dimana pada Al-Qur'an surah Al-Hujurat ayat 11:

اَيَاُيُّهَا الَّذِيْنَ اٰمَنُ. ۖ لَا يَسْخَرْ قَ.وْمٌ ۖمنْ قَ.وْمٍ عَسٰى اَنْ يَّكُوْنُ.وْا خَيْرًا ۖمنْ.هُمْ وَلَا نِسَاۤءٌ ۖمنْ نسَاۤءٍ عَسٰى اَنْ يَّكُنَّ خَيْرًا ۖمنْ.هُنَّ ۚ وَلَا تَ.لْمِزُاوا اَنْ.فُسَكُمْ وَلَا تَ.نَابَ.زُوْا بِالْاَلْقَابِ ۗ بِئْسَ الِاسْمُ الْفُسُوْقُ بَ.عْدَ الْاِيْمَانِ وَمَنْ لَّ يَ.تُبْ فَاُولٰۤئِكَ هُمُ الظّٰلِمُوْنَ

"Wahai orang-orang yang beriman! Janganlah suatu kaum mengolok-olok kaum yang lain (karena) boleh jadi mereka (yang diolok-olok) lebih baik daripada mereka (yang mengolok-olok). Jangan pula perempuan-perempuan (mengolok-olok) perempuan lain (karena) boleh jadi perempuan yang diolok-olok lebih baik daripada perempuan (yang mengolok-olok). Janganlah kamu saling mencela satu sama lain dan janganlah saling memanggil dengan gelar-gelar yang buruk. Seburuk-buruk panggilan adalah (panggilan) yang buruk (fasik) setelah beriman. Barang siapa yang tidak bertobat, maka mereka itulah orang-orang yang zalim." (Q.S Al-Hujurat:11).

Dalam Tafsir Ibnu Katsir menerangkan bahwa orang-orang mukmin adalah saling bersaudara dan sebagai saudara seharusnya saling menjaga dengan tidak mengolok-olok atau merendahkan, kemudian merendahkan atau meremehkan dan mengolok-olok itu diharamkan karena barangkali orang yang diolok-olok dan direndahkan tersebut lebih tinggi kedudukannya disisi Allah SWT, dan lebih disukai oleh-Nya disbanding dengan orang yang merendahkan dan mengolok olok.

Selain itu juga terdapat hadist yang berkaitan dengan topik yang dibahas yakni tentang *cyberbullying*, dimana Al-Bukhari dalam kitab Shahihnya no.6475 dan Muslim dalam kitab Shahihnya no.74 meriwayatkan hadist dari Abu Hurairah bahwa Rasulullah pernah bersabda sebagai berikut.

ومن كان يؤمن بالله واليوم الآخر فليقل خيرا أو ليصمت

"*Dan barangsiapa yang beriman kepada Allah dan hari akhir maka hendaknya dia berkata yang baik atau diam."(HR. Al Bukhari dan Muslim).*

Dalam Tafsir Imam Nawawi pernah membahas tentang hadits ini ketika menjelaskan hadits-hadits Arba'in, bahwa Imam Syafi'i menjelaskan bahwa maksud dari hadits ini adalah apabila seseorang hendak berkata dan kata yang

keluar diperkirakan membawa mudharat maka sebaiknya diam, sedangkan apabila kata yang keluaar tidak membawa mudharat maka silahkan bicara, dan seandainya jika kita yang membelikan kertas untuk para malaikat niscaya kita akan lebih banyak diam dari pada berbicara. Integrasi dari hadits ini terhadap topik yang dibahas pada penelitian adalah pentingnya kita untuk berhati hati dalam menyampaikan kata baik secara lisan maupun dalam teks.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian dan uji coba yang telah dilakukan oleh peneliti, dapat diambil kesimpulan yakni performa tertinggi model didapatkan dengan menerapkan algoritma *Decision Tree* dengan pembobotan TF-IDF dan batas kedalaman maksimum model menggunakan *cross validation* dan *grid search* berhasil mengklasifikasikan teks *cyberbullying* pada data tweet secara optimal menggunakan skenario pembagian data latih dan data uji sebesar 90:10 (6.267 data latih : 696 data uji). Skenario ini mampu menghasilkan performa tertinggi jika dibandingkan dengan skenario 70:30 dan 80:20, dengan rata-rata akurasi sebesar 83,63%, rata-rata Recall 0.812, rata-rata Presisi 0.813, dan rata-rata Skor-F1 0.812. Dengan rata-rata akurasi 83,63% yang ditunjukkan pada model, menunjukkan keberhasilan dan kemampuan model dalam mengklasifikasikan teks *cyberbullying* ke dalam enam kelas yakni religion, age, gender, ethnicity, other cyberbullying, dan not cyberbullying. Meskipun berdasarkan data yang tidak seimbang setelah preproses data teks tweet, model kurang mampu dalam memprediksi teks other cyberbullying dengan not cyberbullying sebab jumlah data pada kedua kelas yang lebih sedikit juga kesamaan kata yang sering muncul dalam dua kelas ini juga membuat model kesulitan melakukan klasifikasi melainkan model cukup mampu memprediksi teks tweet pada kelas label religion, age, gender, dan ethnicity sebab jumlah data yang besar dan juga frekuensi kata yang sering muncul unik.

5.2 Saran

Berdasarkan keterbatasan yang terdapat pada penelitian ini, berikut adalah beberapa saran yang bisa peneliti berikan kepada penelitian berikutnya dengan topik terkait klasifikasi teks ataupun penerapan metode *Decision Tree* pada pengelompokan kata:

1. Pada penelitian teks, keseimbangan jumlah data klasifikasi per kelas merupakan faktor yang penting dan sangat mempengaruhi performa klasifikasi. Oleh sebab itu penelitian berikutnya disarankan untuk memastikan keseimbangan data terlebih dahulu dan apabila data yang digunakan tidak seimbang dapat meng gunakan metode khusus sepserti SMOTE (Synthetic Minority Over-sampling Technique) di mana metode ini digunakan untuk memanipulasi jumlah data pada data *imbalance* untuk meningkatkan nilai performa model secara keseluruhan.
2. Peneliti menyarankan untuk mencoba metode *machine learning* lain yang berbasis pohon keputusan, seperti *Random Forest*, untuk mengklasifikasikan data teks cyberbullying, selain itu juga peneliti berikutnya bisa melakukan perbandingan dari metode tersebut dengan *Decision Tree*, hal ini bertujuan agar dapat diketahui apakah model lain dapat menghasilkan rata-rata akurasi yang lebih tinggi dan lebih stabil dalam menangani kompleksitas teks data tweet jika dibandingkan dengan *Decision Tree*.

DAFTAR PUSTAKA

Fahmy Amin, M. (2022). Confusion Matrix in Binary Classification Problems: A Step-by-Step Tutorial. Journal of Engineering Research, 6(5), 0–0. https://doi.org/10.21608/erjeng.2022.274526

Febriana, T., & Budiarto, A. (2019). Twitter Dataset for Hate Speech and Cyberbullying Detection in Indonesian Language. 2019 International Conference on Information Management and Technology (ICIMTech), 379–382. https://doi.org/10.1109/ICIMTech.2019.8843722

Firmansyah, D., & Suryana, A. (2022). Konsep Pendidikan Akhlak: Kajian Tafsir Surat Al Hujurat Ayat 11-13. Al-Mutharahah: Jurnal Penelitian dan Kajian Sosial Keagamaan, 19(2), 58–82. https://doi.org/10.46781/al-mutharahah.v19i2.538

Imamah, & Rachman, F. H. (2020). Twitter Sentimen Analysis of Covid-19 Using Term Weighting TF-IDF And Logistic Regresion. 2020 6th Information Technology International Seminar (ITIS), 238–242. https://doi.org/10.1109/ITIS50118.2020.9320958

Janisar, A. A., Afzal, H., & Kumar, G. (2021). Identification of HATE speech tweets in Pashto language using Machine Learning techniques. International Journal, 10(3).

Kowalski, R. M., Limber, S. P., & McCatty, R. (2019). Cyberbullying. In S. P. Limber, R. M. Kowalski, & G. W. Giumetti (Eds.), The Wiley Blackwell Handbook of Bullying (pp. 293–310). Wiley-Blackwell.

Margono, H. (n.d.). Analysis of the Indonesian Cyberbullying through Data Mining

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.

Safavian, S. R., & Landgrebe, D. (1991). *A survey of decision tree classifier methodology*. IEEE Transactions on Systems, Man, and Cybernetics, 21(3), 660–674. https://doi.org/10.1109/21.97458

Samalo, H., Pratama, A., & Wijaya, K. (2024). Klasifikasi Cyberbullying pada Media Sosial Menggunakan Algoritma Decision Tree. Jurnal Teknik Informatika dan Sistem Informasi, 10(1), 45-58.

Wang, Y., & Potika, K. (2021). *Cyberbullying detection using machine learning classification methods*. Journal of Information Security and Applications, 58, 102712. https://doi.org/10.1016/j.jisa.2021.102712

LAMPIRAN

Data yang digunakan berdasarkan tiap sentimen

cyberbullying_type	count
religion	7998
age	7992
gender	7973
ethnicity	7961
not_cyberbullying	7945
other_cyberbullying	7823

dtype: int64

Pembuatan dataframe untuk uji coba dan pengolahan data

	text	sentiment
0	In other words #katandandre, your food was cra...	not_cyberbullying
1	Why is #aussietv so white? #MKR #theblock #ImA...	not_cyberbullying
2	@XochitlSuckkks a classy whore? Or more red ve...	not_cyberbullying
3	@Jason_Gio meh. :P thanks for the heads up, b...	not_cyberbullying
4	@RudhoeEnglish This is an ISIS account pretend...	not_cyberbullying

Tahapan proses Cleaning data

	text	sentiment	text_cleaned
0	In other words #katandandre, your food was cra...	5	word katandandr food crapilici mkr
1	Why is #aussietv so white? #MKR #theblock #ImA...	5	aussietv white mkr theblock today sunris studi...
2	@XochitlSuckkks a classy whore? Or more red ve...	5	classi whore red velvet cupcak
3	@Jason_Gio meh. :P thanks for the heads up, b...	5	meh p thank head concern anoth angri dude twitter
4	@RudhoeEnglish This is an ISIS account pretend...	5	isi account pretend kurdish account like islam...

Jumlah data setelah teks cleaning

sentiment	count
0	7946
1	7884
3	7746
5	7637
2	7607
4	5781

Tahapan pengukuran panjang teks

	text	sentiment	text_cleaned	text_len
10	@Jord_Is_Dead http://t.co/UsQInYW5Gn	5		0
6662	@WongDead lameeee	5	lameee	1
6679	@austin_long23 at home .	5	home	1
28994	COINCIDENCE?	4	coincid	1
28928	That's a first.	4	first	1
...	...	...	...	...
45165	@hermdiggz: "@tayyoung_: FUCK OBAMA, dumb ass ...	3	fuck obama dumb ass nigger bitch llthi whore s...	162
44035	You so black and white trying to live like a n...	3	black white tri live like nigger pahahahaha co...	187
30752	I don't retreat.\nyesssssssss http://t.co/Td90k...	4	retreat yesssssssss uh make grownup boruto look...	238
24516	@NICKIMINAJ: #WutKinda\nAt this rate the MKR f...	4	wutkinda rate mkr final decemb mkr haha true u...	352
29205	is feminazi an actual word with a denot...\n@Nas...	4	feminazi actual word denot job mean protect pe...	385

44601 rows × 4 columns

Perbandingan data sebelum dan sesudah EDA

Comparison of Dataset Before and After EDA:

sentiment	count_before_eda	count_after_eda
religion	7998	7893
age	7992	7830
gender	7973	7311
ethnicity	7961	7707
not_cyberbullying	7945	6371
other_cyberbullying	7823	4669
Total	47692	41781

Number of rows removed during EDA: 5911

Detail ringkasan performa skenario berdasarkan metrik dan kedalaman

```
--- Detailed Summary of Metrics Across All Scenarios ---

Scenario 70:30:
  Best Max Depth: None
  Accuracy: 0.8353
  Avg Precision: 0.8117
  Avg Recall: 0.8106
  Avg F1-Score: 0.8106

Scenario 80:20:
  Best Max Depth: None
  Accuracy: 0.8347
  Avg Precision: 0.8102
  Avg Recall: 0.8094
  Avg F1-Score: 0.8095

Scenario 90:10:
  Best Max Depth: None
  Accuracy: 0.8363
  Avg Precision: 0.8133
  Avg Recall: 0.8121
  Avg F1-Score: 0.8121
```

Pengujian overfitting atau underfitting pada model

```
Predictions from Decision Tree for 90:10 split:
First 10 predictions: ['religion', 'gender', 'religion', 'ethnicity', 'religion', 'ethnicity', 'age', 'ethnicity', 'ethnicity', 'religion']
First 10 true labels: ['religion', 'gender', 'religion', 'ethnicity', 'religion', 'ethnicity', 'age', 'ethnicity', 'ethnicity', 'religion']

Model Training Accuracy (from CV): 0.8427
Model Test Accuracy for 90:10 split: 0.8363

Classification Report for 90:10 split:
                    precision    recall  f1-score   support

          religion       0.94      0.91      0.93       790
               age       0.97      0.97      0.97       783
            gender       0.88      0.84      0.86       731
         ethnicity       0.99      0.97      0.98       771
other cyberbullying       0.51      0.58      0.54       467
  not cyberbullying       0.60      0.60      0.60       637

          accuracy                           0.84      4179
         macro avg       0.81      0.81      0.81      4179
      weighted avg       0.84      0.84      0.84      4179
```

Berikut ini merupakan Tahapan Algoritma serta contoh perhitungan dari metode *Decision tree :*

a. Menyusun dataset disusun dari fitur hasil TF-IDF dan label kelas, dapat dilihat Pada Tabel 3.11

   D1, D2, D3, D4 : Kelas 1 (bullying)

D5 : Kelas 0)not bullying)

Dokumen	“fuck” (X1)	“call” (X2)	Label (y)
D1	0.15444	0.12708	1
D2	0	0	1
D3	0	0	1
D4	0	0.46598	1
D5	0	0	0

b. Melakukan perhitungan *Entropy* awal

Total data (S) = 5

Jumlah *bullying* (S_1) = 4, Bukan *bullying* (S_0)

Entropy (Total) = $(-\frac{4}{5} log_2 \frac{4}{5}) + (-\frac{1}{5} log_2 \frac{1}{5})$

Entropy (Total) = $(0.8 * 0.3219) + (0.2 * 2.3219) = 0.7219$

c. Perhitungan *entropry* setiap fitur

1. Fitur “*fuck*” (X1)

Gunakan *threshold* > 0 (artinya kata tersebut muncul)

Muncul (X1 > 0) = D1. Isinya {kelas 1}. Entropy = 0

Tidak muncul (X1 = 0) = D2, D3, D4, D5. Isinya {1,1,1,0}

Entropy = $(-\frac{3}{4} log_2 \frac{3}{4}) + (-\frac{1}{4} log_2 \frac{1}{4}) = 0.8113$

Entropy (S_1 “*fuck*”) = $\frac{1}{5}(0) + \frac{4}{5}(0.8113) = 0.6490$

2. Fitur “*call*” (X2)

Threshold > 0

Muncul (X2 > 0) = D1, D4. Isinya {kelas 1, kelas 1}. *Entropy* = 0

Tidak muncul (X2 = 0) = D2,D3, D5. Isinya {1,1,0}

Entropy = $(-\frac{2}{3} log_2 \frac{2}{3}) + (-\frac{1}{3} log_2 \frac{1}{3}) = 0.9183$

Entropy (S_1 "*call*") = $\frac{2}{5}(0) + \frac{3}{5}(0.9183) = 0.5510$

d. Perhitungan *Information Gain*

- Gain ("*fuck*") = 0.7219 – 0.6490 = 0.0729
- Gain ("*call*") = 0.7219 – 0.5510 = 0.1709

e. Perhitungan *Split Information*

SplitInfo("*fuck*") (Data terbagi 1 dan 4)

$$-(\frac{1}{5} log_2 \frac{1}{5} + \frac{4}{5} log_2 \frac{4}{5}) = 0.7219$$

SplitInfo("*call*") (data terbagi 2 dan 3)

$$-(\frac{2}{5} log_2 \frac{2}{5} + \frac{3}{5} log_2 \frac{3}{5}) = 0.9710$$

f. Perhitungan *Gain Ratio*

Gain Ratio("*fuck*") = 0.0729/0.7219 = 0.1010

Gain Ratio("*call*") = 0.1709/0.9710 = 0.1760

g. Penentuan *node* dan pembentukan pohon

Karena *Gain Ratio* "*call*" (0.1760) lebih besar daripada "*fuck*" (0.1010), maka kata "*call*" dipilih sebagai Node Akar

Struktur pohon :

Apakah "*call*" muncul (TF-IDF > 0) ?

YA = masuk ke kelas 1 (*bullying*)

TIDAK = cek fitur selanjutnya (seperti "fuck")

h. Menentukan Kelas (*Leaf Node*)

Jika pohon dilanjutkan pada cabang "*call* = tidak muncul", maka fitur "*fuck*" akan dihitung kembali. Namun, jika berhenti di sini :

- Daun 1 (*call* > 0) = Dominan kelas 1 (*bullying*)
- Daun 2 (*call* = 0) = Masih ada campuran {1, 1, 0}, maka kelas dominannya adalah 1 (*bullying*)

Aturan keputusan akhir :

IF TF-IDF "*call*" > 0 *THEN Bullying*

ELSE IF TF-IDF "*fuck*" >0 *THEN Bullying*

ELSE Bukan *Bullying* (D5)

Dari hasil uji diatas dapat menunjukkan bahwa dalam algoritma C4.5, kata dengan sebaran nilai TF-IDF yang lebih mampu memisahkan kelas (dalam hal ini "*call*") akan diprioritaskan menjadi *node*

Implementasi tahapan TF-IDF berdasarkan kedalaman dengan sample kata school

```
Original Text: Rebecca Black Drops Out of School Due to Bullying:
Cleaned Text: rebecca black drop school due bulli
Actual Sentiment: not cyberbullying
Model Predicted Sentiment (Full Tree): not cyberbullying

--- Decision Rules for 'school' sample (Full Path) ---
```

1 to 4 of 4 entries Filter

index	Depth	Node Type	Feature	Threshold	Value in Sample	Decision
0	0	Internal	school	0.0413	0.1943	Is 'school' <= 0.0413 (TF-IDF value)? -> No (Go Right)
1	1	Internal	bulli	0.0200	0.1821	Is 'bulli' <= 0.0200 (TF-IDF value)? -> No (Go Right)
2	2	Internal	rebecca	0.5831	0.6062	Is 'rebecca' <= 0.5831 (TF-IDF value)? -> No (Go Right)
3	3	Leaf	Final Prediction	N/A	N/A	Predict: not cyberbullying