BAB 1
PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi telah membawa banyak perubahan signifikan, salah satunya dalam cara berinteraksi dan bertukar informasi. Teknologi yang berkembang ini memperkenalkan platform pesan instan untuk mempermudah interaksi manusia dimanapun berada. Aplikasi pesan instan adalah platform yang menyediakan layanan yang memungkinkan penggunanya untuk mengirimkan pesan teks, video, suara dan berbagai jenis file secara *real-time*. Contoh aplikasi pesan instan yang digunakan saat ini di seluruh dunia adalah Telegram dan Line. Penggunaan aplikasi ini terus bertambah tiap tahunnya, di tahun 2024 tercatat Telegram memiliki 181,57 juta pengguna di Indonesia. Sedangkan Line sendiri cukup popular di wilayah asia, seperti di Jepang dan Indonesia, digunakan oleh 81,27 juta pengguna di Indonesia pada tahun 2024 (The Global Statistics, 2024) . Dengan adanya peningkatan penggunaan Aplikasi pesan instan, tentu saja akan ada tantangan keamanan salah satunya dalam pengumpulan bukti digital yang berbentuk percakapan pesan instan. Dengan populernya penggunaan pesan instan ini, maka bukti digital yang harus diperiksa pun akan lebih banyak dan akan sangat sulit diproses secara manual. Hal ini menciptakan hambatan dalam hal proses investigasi digital, terutama dalam hal mengumpulkan bukti digital. Kejahatan siber seperti penipuan, pemerasan, atau bentuk pelanggaran hukum lainnya sering meninggalkan jejak dalam data percakapan yang dihasilkan oleh aplikasi ini. Tantangan terbesar adalah *volume* data yang sangat besar dan kompleks serta penggunaan enkripsi *end-to-end*, yang membuat pengumpulan dan analisis bukti menjadi sulit (Bogos et al., 2023).

Seiring perkembangan teknologi, bukti digital menjadi sebuah kunci dalam investigasi forensik. Proses pengumpulan bukti digital bergantung pada kemampuan dalam mengakses data yang melewati jaringan. Pengembangan Aplikasi pesan instan tentu saja menggunakan berbagai Teknik dan cara untuk melindungi data yang melewati jaringan. Aplikasi Line dan Telegram menawarkan enkripsi *end-to-end* untuk memastikan keamanan dalam penggunaannya, sehingga hanya dapat diakses oleh pengirim dan penerima. Pada aplikasi Telegram sendiri, enkripsi end-to-end hanya berlaku jika menggunakan fitur *Secret Chats*, sedangkan untuk Line

menerapkan enkripsi lebih menyeluruh (Bogos et al., 2023). Penerapan enkripsi *end-to-end* ini membuat pengumpulan bukti dari aplikasi pesan instan menjadi tantangan tersendiri. Kejahatan siber meninggalkan jejak dari aktivitas kejahatan yang dilakukan sehingga hal ini dapat dijadikan barang bukti. Dalam kasus kejahatan siber barang bukti berupa file dokumen, file *history*, atau file *log* disebut barang bukti digital. Barang bukti pada kasus kejahatan siber berbeda dengan barang bukti kejahatan lainnya, barang bukti tersebut rentan mengalami perubahan atau terkontaminasi sehingga barang bukti tersebut harus disimpan dengan baik di tempat yang aman (Riskiyadi, 2020).

Selain karena popularitas dan peran strategis Line dan Telegram dalam komunikasi digital, kedua aplikasi ini juga dipilih karena memiliki arsitektur penyimpanan data yang mendukung proses forensik digital. Telegram menyimpan sebagian besar data percakapan di *server cloud* kecuali untuk fitur *Secret Chat*, sedangkan Line menyimpan histori *chat* dalam format teks (.txt) di direktori lokal perangkat Android. Arsitektur ini memberikan kemudahan bagi peneliti forensik untuk melakukan akuisisi data tanpa harus melakukan rooting perangkat, sehingga proses ekstraksi menjadi lebih efisien dan *forensically sound* dan tidak merusak integritas data (Apriyani et al., 2023). Selain itu, Line dan Telegram relatif lebih terbuka terhadap proses ekspor data dibandingkan platform lain seperti *WhatsApp* yang memiliki sistem *backup* terenkripsi dan *cloud-key vault* yang tertutup. Telegram, misalnya, menyediakan fitur resmi *ChatExport*, sedangkan Line menyimpan histori pesan yang dapat diakses secara langsung menggunakan perintah *ADB*. Kemudahan ini memberikan peluang besar dalam melakukan penelitian forensik berbasis teks digital (Guna Wicaksana & Gede Suhartana, 2020).

Oleh karena itu dalam proses mengumpulkan, penyimpan dan pengujian bukti digital harus menggunakan standar *National Institute of Justice (NIJ)* agar sah di pengadilan. *National Institute of Justice (NIJ)* adalah sebuah Lembaga yang menetapkan standar untuk forensik digital (Guna Wicaksana & Gede Suhartana, 2020). Aplikasi pesan instan seperti Line dan Telegram menghasilkan *volume* data percakapan yang besar, yang sering kali mengandung informasi penting untuk investigasi. Untuk mendukung pengumpulan bukti digital, metode berbasis *machine learning* dapat digunakan untuk mengidentifikasi pola interaksi yang relevan. Analisis sentimen berbasis kosakata, mampu mendeteksi emosi dan pola dalam percakapan

dengan tingkat akurasi yang tinggi. Dengan memadukan pendekatan ini dengan *Predictive Classification Model (PCM),* data relevan dari aplikasi Line dan Telegram dapat disaring secara lebih efisien.

Salah satu tantangan besar dalam perkembangan komunikasi digital adalah meningkatnya kasus *cyberbullying* atau perundungan siber. *Cyberbullying* merupakan bentuk kekerasan verbal atau psikologis yang dilakukan melalui media digital, seperti pesan teks, komentar, atau percakapan pribadi. Bentuknya bisa berupa penghinaan, pelecehan, intimidasi, atau ujaran kebencian yang ditujukan kepada individu atau kelompok. Perundungan ini kerap terjadi di berbagai platform komunikasi digital, termasuk aplikasi pesan instan, karena sifatnya yang cepat, privat, dan sulit diawasi secara langsung (Aboujaoude et al., 2015). Banyak kasus menunjukkan bahwa pelaku *cyberbullying* memanfaatkan ruang komunikasi pribadi seperti *chat* untuk melakukan tekanan psikologis terhadap korban (Ruslan et al., 2023). Di Indonesia praktik *cyberbullying* dapat dijerat melalui beberapa ketentuan dalam Undang-Undang Informasi dan Transaksi Elektronik (UU ITE), di antaranya Pasal 27 ayat (3) dan ayat (4) UU ITE terkait penghinaan, pencemaran nama baik, serta pemerasan atau pengancaman, serta Pasal 29 UU ITE yang melarang pengiriman informasi elektronik berisi ancaman kekerasan. Selain itu, UU Perlindungan Anak Pasal 76C dan 76D juga mengatur larangan kekerasan psikis terhadap anak, yang mencakup perundungan digital. Hal ini menunjukkan bahwa secara normatif, praktik cyberbullying sudah dapat ditindak melalui mekanisme hukum yang berlaku, meskipun masih bersifat tersebar dan belum terintegrasi dalam satu pasal khusus.

Telegram menjadi salah satu platform yang rawan disalahgunakan karena menyediakan fitur grup publik dan *channel* terbuka. Berdasarkan laporan tahun 2025, sekitar 23% laporan kasus cyberbullying pada pelajar SMP terjadi melalui grup chat tertutup atau terenkripsi seperti Telegram, yang sulit dimoderasi dan minim mekanisme pelaporan. Telegram kerap digunakan untuk menyebarkan ujaran kebencian, *body shaming*, *doxing*, hingga pelecehan seksual. Minimnya moderasi dan lemahnya sistem pelaporan membuat perilaku agresif ini sulit ditindak secara efektif (Negi, n.d.). Sementara itu, aplikasi Line juga banyak digunakan oleh pelajar dan remaja, khususnya di Indonesia dan Jepang. Berbagai laporan media menunjukkan bahwa praktik perundungan digital melalui grup Line seperti pengucilan, penghinaan verbal, hingga penyebaran konten pribadi yang memalukan semakin marak terjadi di

kalangan siswa. Karakter komunikasi yang bersifat privat dan terenkripsi pada Line menyulitkan guru maupun orang tua untuk mendeteksi interaksi berisiko, sehingga banyak kasus baru terungkap setelah korban mengalami dampak psikologis (Nippon.com, 2019). Oleh karena itu, deteksi dini terhadap konten percakapan yang mengandung unsur perundungan menjadi aspek penting dalam investigasi forensik digital, baik untuk pencegahan maupun penindakan hukum (Qibriya et al., 2021). Kedua aplikasi ini dipilih karena selain populer, juga memiliki struktur data yang dapat diakses secara forensik tanpa harus melakukan rooting, serta mewakili dua pendekatan penyimpanan berbeda yaitu cloud dan lokal. Dengan volume percakapan yang tinggi dan penggunaan bahasa informal yang bervariasi, pendekatan manual menjadi tidak efisien. Diperlukan sistem otomatis yang mampu mengklasifikasikan konten berisiko tinggi secara akurat dan cepat, salah satunya dengan memanfaatkan metode *machine learning* seperti *Predictive Classification Model (PCM)* (Lee et al., 2023).

Pendekatan yang dapat digunakan adalah *Predictive Classification Model (PCM). Predictive Classification Model (PCM)* adalah sebuah pendekatan yang memungkinkan untuk memprediksi dan mengklasifikasikan data-data digital berdasarkan pola perilaku pengguna. Dengan *Predictive Classification Model (PCM),* analisis pola perilaku pengguna Telegram dan line dapat dilakukan dengan memanfaatkan data yang tersedia, seperti *log* aktivitas tentang kapan dan dengan siapa pengguna berinteraksi. Penggunaan *Predictive Classification Model (PCM)* pada aplikasi Telegram dan Line menjadi sebuah Solusi untuk menghadapi keterbatasan akses ke data dikarenakan enkripsi *end-to-end*. Selain itu, perbandingan performansi *Predictive Classification Model (PCM),* dalam pengumpulan bukti digital pada aplikasi Line dan Telegram menjadi penting. Perbedaan struktur data dan penerapan enkripsi di kedua aplikasi mempengaruhi efisiensi dan akurasi model *Predictive Classification Model (PCM),* dalam menyaring dan menganalisis data (Aziz et al., 2018). Untuk mendukung proses klasifikasi otomatis dalam *Predictive Classification Model (PCM),* penelitian ini menggunakan dua algoritma machine learning yang mewakili pendekatan berbeda, yaitu *Complement Naive Bayes* dan *Random Forest*. *Complement Naive Bayes* dipilih karena kemampuannya dalam menangani data teks yang tidak seimbang serta efisiensinya dalam mengidentifikasi kolom dan konten percakapan pada aplikasi pesan instan. Model ini dapat digunakan untuk memfilter

data yang tidak relevan secara otomatis dan efektif, sehingga sesuai untuk diterapkan dalam kerangka investigasi digital yang membutuhkan kecepatan dan presisi pada data berbasis percakapan (Lee et al., 2023). Sementara itu, *Random Forest* digunakan karena kekuatannya dalam mengolah relasi kompleks antar fitur dan kestabilannya dalam melakukan klasifikasi menggunakan pendekatan berbasis voting dari sejumlah pohon keputusan. Algoritma ini digunakan untuk mendeteksi kejahatan siber pada media sosial, dengan pendekatan yang memanfaatkan parameter-parameter seperti informasi pengguna dan ciri linguistik dalam komentar. *Random Forest* terbukti mampu mengenali pola-pola yang menunjukkan ancaman atau aktivitas mencurigakan dalam data teks publik, serta bersifat robust terhadap data berisik dan bervariasi, sehingga sangat relevan untuk kebutuhan analisis forensik digital (Arora et al., 2019). Kedua model diterapkan secara terpisah untuk keperluan perbandingan performa dalam mendeteksi komentar *cyberbullying*. Pendekatan ini dipilih agar dapat memberikan gambaran menyeluruh mengenai efektivitas masing-masing algoritma dalam menangani variasi karakteristik data percakapan, serta mendukung pengambilan keputusan dalam pemilihan model klasifikasi terbaik yang sesuai dengan kerangka kerja investigasi digital berbasis standar *National Institute of Justice (NIJ).*

Oleh karena itu, penelitian ini tidak hanya berfokus pada proses ekstraksi dan pengolahan bukti digital dari aplikasi Line dan Telegram, tetapi juga membandingkan performa dua algoritma *machine learning*, yaitu *Complement Naive Bayes* dan *Random Forest*, yang masing-masing diimplementasikan secara terpisah dalam metode *Predictive Classification Model (PCM).* Perbandingan ini bertujuan untuk menentukan model yang paling efektif dalam mengklasifikasikan komentar berbasis teks dalam konteks investigasi forensik digital terhadap kasus *cyberbullying*. Dengan demikian, penggunaan standar *National Institute of Justice (NIJ)* dalam memastikan validitas bukti digital dan *Predictive Classification Model (PCM)* dalam analisis data, dapat menjadi sebuah solusi efektif dalam menghadapi tantangan dalam pengumpulan bukti digital.

1.2 Rumusan Masalah

Rumusan masalah penelitian ini adalah, sebagai berikut:

1. Bagaimana proses metadata bukti digital dari aplikasi Line dan Telegram berdasarkan standar *National Institute of Justice (NIJ)* untuk digunakan sebagai bukti digital?

2. Bagaimana metode *Predictive Classification Model (PCM)* melakukan pemeriksaan dan klasifikasi data digital hasil akuisisi dalam mendukung investigasi forensik digital menggunakan algoritma *Complement Naive Bayes* dan *Random Forest*?

3. Bagaimana perbandingan performansi metode *Predictive Classification Model (PCM)* menggunakan algoritma *Complement Naive Bayes* dan *Random Forest* dalam mengklasifikasikan bukti digital?

1.3 Batas Penelitian

Penelitian ini memiliki beberapa batasan untuk memastikan tetap terfokus dan relevan, antara lain:

1. Data yang dikumpulkan secara forensik dalam penelitian ini terbatas pada dua aplikasi pesan instan, yaitu Line dan Telegram.
2. Fokus penelitian adalah pada aspek pengumpulan bukti digital dan analisis forensik.
3. Bukti digital yang dikumpulkan dalam penelitian ini hanya berupa *log* aktivitas, riwayat pesan, metadata, serta data yang relevan dalam konteks investigasi forensik.
4. Predictive *Classification Model (PCM)* hanya digunakan untuk menguji proses klasifikasi pada dataset terpisah yang berasal dari platform Kaggle, bukan pada data hasil ekstraksi dari Line dan Telegram.

1.4 Tujuan

Penelitian ini bertujuan :

1. Mengidentifikasi dan menganalisis proses pengumpulan metadata bukti digital dari aplikasi Line dan Telegram dengan mengacu pada standar *National Institute of Justice (NIJ)*, guna menjamin validitas dan legalitas bukti digital dalam proses investigasi forensik.
2. Menerapkan metode *Predictive Classification Model (PCM)* dalam proses pemeriksaan dan klasifikasi terhadap data digital hasil akuisisi, dengan menggunakan dua algoritma *machine learning* yang berbeda, yaitu *Complement Naive Bayes* dan *Random Forest*, untuk mendukung efektivitas investigasi forensik digital terhadap kasus *cyberbullying*.

3. Mengevaluasi dan membandingkan performa dua algoritma *machine learning*, yaitu *Complement Naive Bayes* dan *Random Forest*, dalam penerapan *Predictive Classification Model (PCM)* untuk klasifikasi otomatis terhadap bukti digital berbasis teks.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat, antara lain:

1. Mengembangkan model prediksi berbasis *machine learning* melalui *Predictive Classification Model (PCM)* yang dapat membantu dalam pengumpulan bukti digital.
2. Mempermudah dalam mengumpulkan bukti digital yang valid dan dapat dipertanggungjawabkan dengan standar *National Institute of Justice (NIJ).*
3. Meningkatkan efisiensi dan akurasi proses investigasi digital, khususnya dalam kasus-kasus yang melibatkan aplikasi Line dan Telegram.