# ANALISIS FAKTOR RISIKO DAN PREDIKSI DATA BREACH MENGGUNAKAN ALGORITMA RANDOM FOREST BERBASIS VERIZON COMMUNITY DATABASE (VCDB)

**Nasrullah Akbar Fadriansyah[1], Yudi Alfarizi[2], Achmad Syahrudin[3], Aditya April Riandi[3], Albert Riyandi[3]**
nasrullahakbar18@gmail.com[1] , yudi.alfaarizi@gmail.com[2], achmad.asy@gmail.com[3], aditya.apriliandi10@gmail.com[3], albert.abe@bsi.ac.id[3]
**Universitas Bina Sarana Informatika**

**ABSTRAK**

Di era transformasi digital, ancaman keamanan siber, khususnya insiden pelanggaran data (data breach), berkembang semakin canggih melalui skenario serangan multi-tahap yang mengeksploitasi kerentanan teknis maupun anomali perilaku manusia. Pendekatan keamanan tradisional yang bersifat reaktif terbukti tidak lagi memadai untuk melindungi aset digital secara komprehensif. Penelitian ini bertujuan untuk menganalisis secara mendalam faktor-faktor risiko fundamental yang berkontribusi terhadap keberhasilan data breach sekaligus mengusulkan kerangka prediksi proaktif berbasis Machine Learning. Dengan memanfaatkan Verizon Community Database (VCDB) sebagai landasan data intelijen historis global, penelitian ini mengimplementasikan algoritma Random Forest untuk mengidentifikasi dan memprediksi anomali serangan siber. Penelitian ini mengolah total data awal sebanyak 10.586 insiden, yang kemudian menyisakan 10.364 data setelah melalui tahap cleansing. Dataset tersebut dibagi secara terstratifikasi menjadi 8.291 data untuk training (80%) dan 2.073 data untuk testing (20%). Mengingat adanya ketidakseimbangan kelas target yang ekstrem, teknik Synthetic Minority Over-sampling Technique (SMOTE) diterapkan pada data training sehingga menghasilkan 17.955 sampel seimbang. Hasil penelitian menunjukkan bahwa fitur data pribadi (Personally Identifiable Information) dan data rekam medis menjadi prediktor dengan tingkat kepentingan (feature importance) tertinggi dalam memicu kebocoran informasi. Model Random Forest Classifier yang dibangun berhasil mencapai tingkat akurasi akhir sebesar 83,12% (0,8312) dalam mengklasifikasikan status ancaman ke dalam tiga kategori (Data Breach, No Breach, dan Potential Breach). Pemodelan prediktif ini berkontribusi dalam menghadirkan sistem kecerdasan ancaman (threat intelligence) yang otomatis dan presisi, memungkinkan organisasi untuk merancang strategi mitigasi preventif sebelum intrusi siber berakibat fatal

**Kata Kunci:** Mahasiswa, Perilaku Off-Task, Single Subject Design, Token Economy.

## PENDAHULUAN

Di era digital yang dibanjiri oleh kelimpahan data, organisasi di seluruh dunia terus mencari cara baru untuk membuat, menyimpan, dan mentransfer informasi dengan kecepatan yang terus meningkat. Ketergantungan yang tinggi pada data ini membuat praktik tata kelola dan keamanan data—baik dari ancaman internal maupun eksternal—menjadi prioritas paling krusial bagi setiap institusi. Transformasi sistem kritis yang kini banyak didukung oleh teknologi cerdas menuntut agar keamanan informasi tidak lagi dipandang sebelah mata, melainkan harus diintegrasikan dengan pemahaman tentang kerentanan sistem terhadap berbagai vektor serangan baru (Roy et al., 2023).

Seiring berjalannya waktu, ancaman siber telah tumbuh menjadi jauh lebih canggih dan sering terjadi. Pendekatan keamanan siber tradisional yang bersifat reaktif telah terbukti tidak lagi memadai untuk melindungi aset digital dan infrastruktur kritis secara komprehensif (Mareedu, 2023). Sebagai respons terhadap keterbatasan ini, arsitektur keamanan siber modern harus bertransisi menuju model prediktif berbasis data yang didukung oleh kecerdasan ancaman (threat intelligence) waktu-nyata untuk mendeteksi

indikator kompromi lebih awal.

Namun, sebagian besar sistem deteksi saat ini masih berkutat pada skala waktu jangka pendek, seperti memindai log sistem dan lalu lintas jaringan untuk mencari tanda bahaya saat serangan sedang atau baru saja terjadi. Sangat sedikit upaya yang dikerahkan untuk melakukan prediksi serangan siber dalam jangka panjang secara holistik. Padahal, pendekatan proaktif yang mampu meramalkan tren serangan sebelum terjadi sangat diinginkan karena dapat memberikan waktu yang cukup bagi para pemangku kepentingan untuk mengembangkan dan mendistribusikan alat pertahanan yang tepat (Almahmoud et al., 2023).

Untuk mewujudkan sistem pertahanan yang proaktif tersebut, pemanfaatan kecerdasan buatan, khususnya Machine Learning (ML), telah menjadi landasan utama. Berbagai penelitian telah mengkaji efektivitas model ML dalam melakukan klasifikasi dan deteksi berbagai ancaman siber, mulai dari rekayasa sosial (phishing) hingga penyusupan jaringan (malware dan serangan DDoS) (Ahsan et al., 2022). Model-model ini mampu memitigasi risiko dengan cara mengidentifikasi pola kompleks dari data historis yang sering kali luput dari pengawasan analisis manusia konvensional.

Digitalisasi dan revolusi Internet of Things (IoT) menghasilkan data keamanan siber dalam volume yang sangat masif, sehingga analisis intelijen data dan otomatisasi menjadi suatu keharusan. Memanfaatkan pengetahuan kecerdasan buatan terbukti esensial dalam menyediakan sistem keamanan yang dinamis, otomatis, dan mutakhir. Ekstraksi wawasan berharga dari data siber secara cerdas memungkinkan pengambil keputusan untuk merancang perlindungan generasi berikutnya yang jauh lebih efisien dalam menyelesaikan anomali siber (Sarker, 2022).

Lebih spesifik lagi, penerapan Machine Learning terus mengalami kemajuan pesat dalam domain prediksi kejahatan siber (cybercrime). Algoritma-algoritma mutakhir kini tidak hanya digunakan untuk deteksi intrusi dasar, tetapi juga dirancang untuk memprediksi perilaku kriminal di dunia maya dengan mendeteksi anomali perilaku eksploitatif pada tahap awal. Penyesuaian algoritma pembelajaran mesin telah membuka peluang riset yang signifikan untuk menambal celah keamanan organisasi sebelum insiden fatal terjadi (Elluri et al., 2023).

Tantangan terbesar dalam prediksi kejahatan siber saat ini adalah menghadapi serangan tingkat lanjut yang memiliki banyak fase (multi-tahap), yang sering disamarkan di balik aktivitas sistem yang tampak normal. Serangan zero-day dan Advanced Persistent Threats (APT) mengharuskan sistem deteksi memiliki tingkat akurasi tinggi yang tidak mudah dikelabui. Pendekatan prediktif yang menggabungkan peningkatan algoritma komputasi klasifikasi tinggi terbukti mampu mencapai tingkat akurasi yang lebih superior dalam mendeteksi skenario serangan multi-tahap ini (Dalal, Manoharan, et al., 2023).

Di sisi lain, kelalaian maupun anomali perilaku dari entitas manusia sering kali menjadi titik lemah utama yang memfasilitasi terjadinya pelanggaran data (data breach). Oleh karena itu, pengawasan tidak boleh hanya terbatas pada infrastruktur teknis, melainkan juga harus mencakup pemrosesan log aktivitas pengguna secara real-time. Pembuatan profil dari jejak digital pengguna yang menggabungkan framework deteksi anomali perilaku (User Behavior Anomaly Detection) sangat krusial untuk mencegah serangan umum seperti ransomware dan pencurian kredensial (Folino et al., 2023).

Selain perilaku pengguna tingkat individu, karakteristik struktural dan operasional pihak ketiga dalam rantai pasokan digital sebuah institusi juga memberikan dampak langsung terhadap kerentanan keamanan. Data empiris berskala besar membuktikan bahwa atribut jaringan dan karakteristik operasional tersebut merupakan prediktor yang sangat

signifikan untuk menilai risiko terjadinya serangan siber berupa data breach (Hu et al., 2023). Pemanfaatan data historis serangan pihak ketiga terbukti mampu menambahkan daya deteksi substansial untuk menilai seberapa rentan sebuah entitas perusahaan.

Untuk membangun kerangka prediksi serangan siber generasi berikutnya yang tangguh, pendekatan algoritma berbasis Decision Tree yang dioptimalkan dan di-kombinasikan dengan metode pengklasifikasian multi-kelas telah diakui keandalannya dalam menangani data trafik ancaman yang sangat besar dan kompleks. (Dalal, Lilhore, et al., 2023)Berangkat dari kesenjangan penelitian mengenai pentingnya faktor prediktif serta keandalan machine learning berbasis tree, penelitian ini bertujuan untuk menguji kemampuan algoritma Random Forest—sebagai salah satu model ensemble berbasis pohon yang paling kuat dalam menangkal overfitting—untuk memprediksi insiden data breach. Sebagai landasan utama, penelitian ini menggunakan Verizon Community Database (VCDB), sebuah repositori data historis insiden siber global komprehensif, untuk menganalisis secara mendalam faktor-faktor risiko mana yang paling berkontribusi terhadap keberhasilan suatu pelanggaran data di berbagai sektor industri.

Melengkapi analisis kerentanan sistem secara global, penting untuk meninjau bahwa risiko data breach tidak hanya bergantung pada infrastruktur teknis, tetapi juga sangat dipengaruhi oleh dinamika organisasional di berbagai sektor spesifik yang menyimpan data bernilai tinggi, seperti institusi pendidikan dan riset. Lingkungan organisasi dengan keterbukaan kolaboratif yang tinggi cenderung memperluas jejak digital pengguna yang dapat dengan mudah dieksploitasi oleh peretas. Namun demikian, membatasi aliran data secara domestik terbukti tidak serta-merta menjamin keamanan arsitektur jaringan; sebaliknya, entitas yang memiliki volume arus data lintas batas (cross-border data flow) yang tinggi justru sering menunjukkan tingkat pelaporan insiden data breach yang lebih rendah karena mereka diwajibkan untuk mematuhi regulasi tata kelola yang lebih ketat serta memiliki tingkat kesadaran manajemen perlindungan data yang lebih matang (Li et al., 2023).

Sebagai respons terhadap tingginya kompleksitas pengelolaan data tersebut, adopsi penyimpanan komputasi awan (cloud storage) muncul sebagai benteng pelindung teknis yang sangat krusial bagi kelangsungan keamanan sebuah organisasi. Berlawanan dengan asumsi konvensional yang kerap menganggap basis data on-premise lebih terlindungi, penggunaan infrastruktur cloud—khususnya public cloud—terbukti secara signifikan mampu mengurangi dampak kerugian dari insiden pelanggaran data, mengingat penyedia layanan mampu mengintegrasikan standar perlindungan keamanan eksternal yang jauh lebih mutakhir dan terpusat. Di sisi lain, peningkatan frekuensi pengungkapan celah keamanan secara publik (vulnerability disclosure) justru berpotensi memperbesar peluang eksploitasi oleh penyerang oportunistik, yang secara tidak langsung semakin menggarisbawahi besarnya urgensi pemanfaatan arsitektur cloud untuk memitigasi kerentanan perangkat lunak tersebut secara terpadu (Li et al., 2023).

Meskipun infrastruktur komputasi awan menawarkan skalabilitas yang sangat baik, sistem yang sangat mengandalkan pembagian sumber daya fisik (shared resources) seperti Mesin Virtual (VM) mau tidak mau juga membuka jalan bagi serangkaian vektor serangan siber yang sama sekali baru. Kesalahan konfigurasi alokasi ruang dan celah kerentanan pada tingkat hypervisor dapat dimanfaatkan secara cerdik oleh pengguna jahat untuk menyusup dan mencapai posisi ko-residensi (co-residency) dengan sistem target mereka, yang sering kali berujung pada eksploitasi pencurian data sensitif secara kaskade melalui jaringan (network-cascading effect). Eskalasi ancaman semacam ini pada akhirnya menuntut adanya sebuah mekanisme pertahanan berbasis sistem yang sanggup memonitor serta

mengkuantifikasi status keamanan VM dari berbagai metrik risiko sebelum intrusi siber menyebar melintasi batas fisik penyedia layanan(Saxena et al., 2023).

Untuk melindungi kerahasiaan data komputasional dari ancaman kebocoran akibat eksploitasi co-residency tersebut, pendekatan deteksi tradisional yang hanya mengandalkan pemisahan paksa sumber daya fisik terbukti menjadi metode yang usang dan sangat menguras biaya operasional. Oleh karena itu, penerapan model prediksi ancaman cerdas yang digerakkan oleh algoritma Machine Learning, seperti arsitektur Extreme Gradient Boosting (XGB), menjadi esensial untuk memproses rekam jejak perilaku akses secara real-time dan mengekstrak korelasi dari pola serangan historis. Melalui integrasi pemodelan yang mengevaluasi multi-faktor risiko secara dinamis, sistem pengambil keputusan dapat secara proaktif mengantisipasi aktivitas lalu lintas jaringan yang mencurigakan, sehingga manuver penanganan seperti migrasi data antar VM dapat dieksekusi sedini mungkin sebelum pelanggaran sistem berhasil dilakukan penyerang (Saxena et al., 2023).

Selain kompleksitas arsitektur awan, hambatan terbesar dalam memproses dan memprediksi anomali serangan secara presisi pada lalu lintas Internet of Things (IoT) maupun basis data insiden raksasa adalah karakteristik log yang acap kali didominasi oleh kelas data tidak seimbang (imbalanced data). Kondisi empiris di mana volume rekam jejak aktivitas jaringan normal jauh lebih melimpah dibandingkan dengan sampel anomali serangan menyebabkan sebagian besar model analitik rentan mengalami fenomena underfitting dan bias prediksi terhadap kelas mayoritas. Jika tidak ditangani secara khusus melalui pra-pemrosesan data, ketimpangan kelas ini akan menyebabkan model gagal mengenali ancaman siber sporadis dan mendistorsi metrik akurasi evaluasi secara drastis saat menghadapi Distributed Denial of Service (DDoS) maupun teknik pencurian data lainnya (Alkhudaydi et al., 2023).

Untuk mengeliminasi dampak ketimpangan distribusi fitur pada log keamanan jaringan tersebut, optimasi teknik pra-pemrosesan seperti Synthetic Minority Over-sampling Technique (SMOTE) dapat dipadukan dengan sangat efektif menggunakan arsitektur klasifikasi ensemble tingkat lanjut. Dengan menyeimbangkan komposisi dataset melalui interpolasi linear untuk menciptakan duplikasi sampel kelas serangan minoritas secara sintetik, varian algoritma ansambel berbasis pohon keputusan, seperti Random Forest dan XGBoost, terbukti mampu menekan bias algoritma dan mengatrol tingkat kemampuan deteksi secara ekstrem hingga menyentuh rentang akurasi di atas 98%. Validasi empiris pada optimalisasi kelas minoritas ini semakin mempertegas konklusi bahwa kombinasi antara rekayasa penyeimbangan data dan algoritma pohon klasifikasi merupakan metodologi terbaik untuk memetakan variasi trafik ancaman dengan presisi yang tangguh (Alkhudaydi et al., 2023).

Di sisi lain, secanggih apa pun algoritmanya dalam mengelola data kategorikal berdimensi tinggi, arsitektur murni dari Random Forest tetap akan membentur batasan tertentu ketika dihadapkan pada skema ancaman siber multi-tahap tingkat lanjut yang keterikatannya berpusat pada dependensi urutan waktu (sekuensial). Mengingat peretas dewasa ini terus berevolusi menyebarkan injeksi melalui anomali jaringan temporal yang sangat dinamis, penggabungan model klasifikasi statis ini dengan arsitektur jaringan saraf tiruan layaknya Long Short-Term Memory (LSTM) dipercaya mampu menghadirkan sebuah kerangka kerja hibrida (hybrid) yang komprehensif. Di dalam kolaborasi pemrosesan arsitektural ini, mesin Random Forest akan dialokasikan sebagai filter awal untuk mengklasifikasikan pola struktural statis—seperti abnormalitas ukuran paket atau upaya akses basis data historis—secara gesit dan hemat daya komputasi (Hammad, 2024).

Sebagai kelanjutannya, pendelegasian pengawasan anomali dinamis kepada memori jangka panjang jaringan LSTM terbukti amat superior untuk menyaring sisa-sisa ancaman manipulasi lalu lintas persisten yang bergerak luput melalui rentang waktu tertentu. Orkestrasi pengenalan pattern dari kedua fondasi kecerdasan buatan ini tidak hanya menggenjot metrik evaluasi akurasi deteksi ancaman gabungan secara signifikan hingga menembus angka 97%, namun juga berperan krusial dalam menekan serendah mungkin persentase negatif palsu (false negatives). Keseimbangan komplementer antara ketajaman presisi logika klasifikasi pohon keputusan dan tingkat sensitivitas historis dari arsitektur deep learning tersebut akan memastikan bahwa insiden pelanggaran data yang disamarkan sehalus mungkin tetap akan memicu alarm sistem mitigasi otomatis (Hammad, 2024).

Meskipun demikian, ambisi untuk memperkuat kapasitas prediktif semua model machine learning tersebut dengan melatihnya menggunakan basis log global kerap kali tersendat oleh ketatnya kepatuhan regulasi privasi saat mencoba mengumpulkan data sensitif ke dalam satu server pemrosesan terpusat. Untuk mengurai kebuntuan struktural ini, paradigma Federated Learning (Pembelajaran Terfederasi) menawarkan mekanisme revolusioner di mana pelatihan pembaruan node-node klasifikasi pada algoritma Random Forest dapat ditransfer serta didistribusikan langsung secara mandiri pada perangkat sistem klien atau organisasi partisipan. Melalui skema kolaboratif desentralisasi ini, seluruh entitas keamanan di berbagai sektor industri dapat saling bertukar pembaruan parameter statistik terkait profil ancaman zero-day terbaru, tanpa harus mengekspos dan menyetorkan wujud data mentah lalu lintas internal mereka yang sangat rawan memicu kebocoran (data breach) (Markovic et al., 2024).

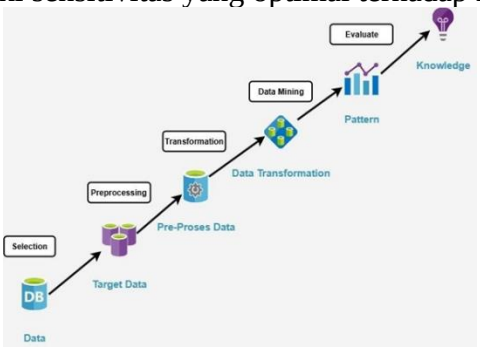
Sebagai penyempurnaan dari protokol pelindungan model tersebut, fakta bahwa pertukaran parameter internal hasil pelatihan yang dilempar kembali ke server agregator masih menyisakan celah bagi peretas ahli untuk merangkai ulang struktur informasi aslinya (reconstruction attacks) memicu diterapkannya enkripsi tingkat lanjut. Integrasi mekanisme Differential Privacy langsung pada jantung fondasi pembangunan pohon Random Forest—yakni dengan sengaja menginjeksikan desibel noise matematis acak pada kalkulasi akhir daun klasifikasi—diklaim mampu mendirikan jaring pengaman anonimitas yang definitif bagi seluruh rekaman aktivitas user. Sekalipun induksi kerahasiaan berlapis ini secara harfiah akan mereduksi secuil efisiensi absolut klasifikasi, penggunaan arsitektur tersebut menjadi afirmasi tak terbantahkan bahwa sistem analitik pendeteksi data breach yang menambang basis Verizon Community Database (VCDB) ini bersifat privat secara default dan mutlak mustahil disalahgunakan oleh pihak ketiga sebagai perangkat eksploitasi baru (Markovic et al., 2024).

## METODE

Penelitian ini menggunakan pendekatan kuantitatif eksplanatori yang bertujuan untuk menjelaskan faktor-faktor risiko yang memengaruhi probabilitas terjadinya data breach pada berbagai entitas organisasi. Pendekatan ini dipilih sebab memungkinkan pengujian hubungan antar variabel secara empiris melalui analisis data rekam jejak ancaman siber berskala besar, dan mendukung pembangunan model prediktif berbasis machine learning. Populasi dalam penelitian ini ialah seluruh entri insiden keamanan siber yang tercatat dalam Verizon Community Database (VCDB). Dataset ini mencakup puluhan ribu insiden dari berbagai negara dengan latar belakang sektor industri dan karakteristik teknis yang beragam. Seluruh responden (insiden) yang memiliki data lengkap dan relevan terhadap variabel penelitian digunakan sebagai sampel penelitian, sehingga penelitian ini bersifat census-based tanpa proses pengambilan sampel tambahan.

Variabel dependen pada penelitian ini ialah status data breach, yang di klasifikasikan ke dalam tiga kategori: Data Breach, No Breach, dan Potential Breach. Variabel independen meliputi faktor risiko teknis dan operasional, seperti jenis aktor penyerang , vektor aksi, pola serangan , aset yang menjadi target, jenis data sensitif yang diincar

Analisis data dilakukan menggunakan pendekatan Knowledge Discovery in Database (KDD) yang meliputi tahapan seleksi data, preprocessing, transformasi data, pemodelan, dan interpretasi hasil. Data sekunder dalam format CSV diolah memakai bahasa pemrograman Python menggunakan pustaka pandas, numpy, scikit-learn, dan imblearn. Proses preprocessing mencakup pembersihan data (cleansing), pengkodean variabel kategorikal (one-hot encoding), dan penanganan ketidakseimbangan kelas (imbalanced data) menggunakan teknik Synthetic Minority Over-sampling Technique (SMOTE) untuk memastikan model memiliki sensitivitas yang optimal terhadap kelas minoritas.

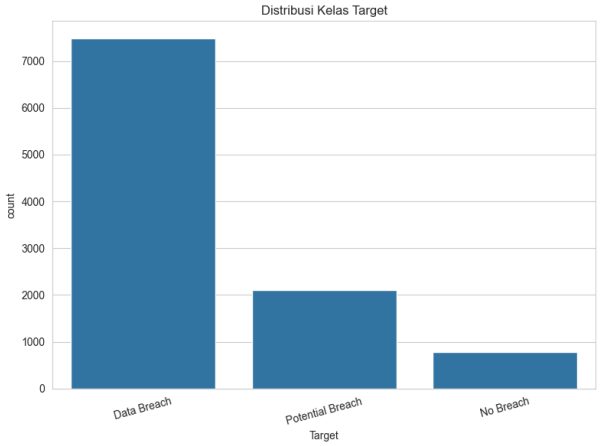



Gambar 1. Model konseptual proses Knowledge Discovery in Database (KDD) yang digunakan dalam penelitian ini.

Pemodelan dilakukan menggunakan algoritma Random Forest Classifier untuk memprediksi tingkat risiko data breach serta mengidentifikasi faktor-faktor yang paling berpengaruh melalui analisis feature importance. Kinerja model dinilai menggunakan metrik akurasi, precision, recall, F1-score, dan confusion matrix, dan divalidasi menggunakan teknik k-fold cross-validation untuk memastikan keandalan hasil prediksi terhadap data insiden siber yang baru.

## HASIL DAN PEMBAHASAN

**Evaluasi Model** Penelitian ini menerapkan algoritma *Random Forest Classifier* untuk mengklasifikasikan status ancaman dan insiden keamanan siber ke dalam tiga kategori utama, yaitu *Data Breach* (pelanggaran berhasil), *No Breach* (serangan digagalkan), dan *Potential Breach* (indikasi kompromi). Pemilihan *Random Forest* didasarkan pada karakteristik data log keamanan siber dari Verizon Community Database (VCDB) yang berskala masif, bersifat sangat heterogen, serta mengandung interaksi non-linear yang kompleks antar variabel serangan. Dalam konteks penelitian keamanan informasi, model klasifikasi yang mampu menangkap korelasi rumit antara vektor serangan, jenis aset, dan profil korban menjadi sangat krusial agar hasil prediksi intelijen ancaman (*threat intelligence*) tidak bersifat reduksionis. Oleh karena itu, *Random Forest* dipandang sebagai pendekatan yang paling sinkron dibandingkan algoritma linier konvensional.




**Grafik 1. Distribusi Kelas Target pada Dataset VCDB**

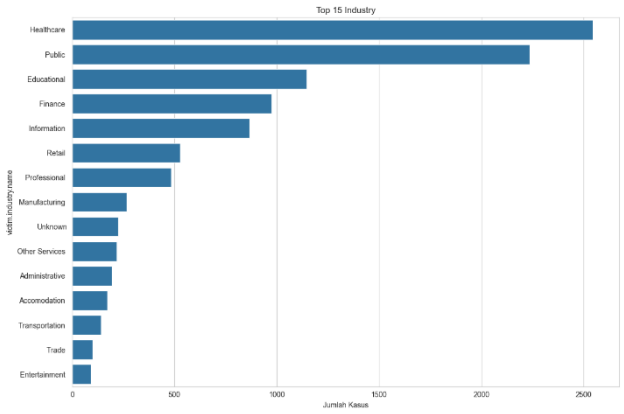
Tahap awal yang sangat krusial sebelum melangkah pada evaluasi pemodelan prediktif adalah memahami karakteristik dan struktur fundamental dari data historis yang diolah. Berdasarkan Grafik 1, terlihat jelas adanya fenomena distribusi kelas target yang sangat tidak seimbang (imbalanced data). Dalam visualisasi tersebut, insiden yang pada akhirnya gagal dibendung dan berujung pada Data Breach (pelanggaran data yang berhasil) mendominasi secara drastis dengan jumlah melampaui angka 7.000 kasus. Volume ini terpaut sangat jauh dan asimetris jika dibandingkan dengan kelas Potential Breach (indikasi atau potensi kompromi) yang hanya berada di kisaran 2.000 kasus, serta kelas No Breach (serangan yang berhasil dideteksi dan digagalkan) yang jumlahnya bahkan kurang dari 1.000 kasus.

Tingginya frekuensi pelaporan pada kelas Data Breach ini mencerminkan realitas empiris dalam industri keamanan siber dan pelaporan insiden global. Institusi dan organisasi cenderung lebih komprehensif dan diwajibkan secara regulasi untuk mendokumentasikan insiden yang telah menimbulkan kerugian nyata atau kebocoran informasi sensitif. Sebaliknya, ribuan serangan siber yang berhasil dipatahkan oleh sistem keamanan perimeter jaringan (seperti firewall atau Intrusion Prevention System) setiap harinya sering kali dianggap sebagai anomali lalu lintas biasa dan tidak tercatat secara mendetail di dalam repositori publik seperti VCDB.

Dari perspektif pembelajaran mesin (machine learning), ketimpangan kelas yang ekstrem ini menghadirkan masalah algoritmik yang serius. Jika dataset ini langsung digunakan untuk melatih algoritma Random Forest tanpa intervensi data, model akan mengalami fenomena yang disebut sebagai bias kelas mayoritas (majority class bias) atau underfitting pada kelas minoritas. Mesin analitik akan cenderung mengejar nilai akurasi semu (accuracy paradox) dengan cara terus-menerus memprediksi setiap insiden sebagai Data Breach untuk mendapatkan skor tinggi, namun akan gagal total dan buta dalam mengenali pola anomali dari serangan yang sebenarnya bisa ditahan (No Breach).

Oleh karena itu, kondisi ketimpangan struktural ini menuntut adanya intervensi matematis melalui penerapan Synthetic Minority Over-sampling Technique (SMOTE) pada tahap pra-pemrosesan (preprocessing). Berbeda dengan teknik random oversampling konvensional yang sekadar menduplikasi baris data yang sudah ada, SMOTE bekerja dengan cerdas menggunakan pendekatan k-Nearest Neighbors (k-NN) untuk menginterpolasi fitur dari sampel kelas minoritas yang berdekatan, guna mensintesis data artifisial baru yang unik di sepanjang ruang fitur linear. Melalui proses rekayasa penyeimbangan ini, ruang sampel minoritas diperbesar hingga setara dengan kelas mayoritas, sehingga algoritma dipaksa untuk mempelajari seluruh spektrum skenario

ancaman siber. Hasilnya, model Random Forest yang dibangun tidak akan bias dan mampu memiliki tingkat sensitivitas (recall) yang adil serta tangguh untuk setiap kategori klasifikasi serangan.




**Grafik 2. Distribusi 15 Industri Teratas Korban Serangan Siber**

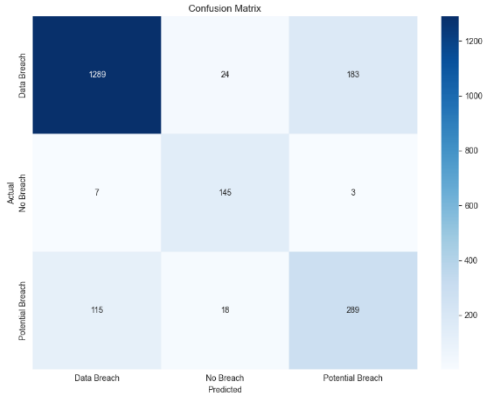
Selain mengevaluasi keseimbangan kelas target, keandalan sebuah model prediktif juga harus dipahami dan dikalibrasi dalam konteks demografi korban. Peta persebaran korban ini sangat penting untuk memahami dari mana asal data log yang dipelajari oleh algoritma. Grafik 2 memperlihatkan distribusi 15 industri teratas yang paling sering menjadi sasaran eksploitasi siber. Sektor kesehatan (Healthcare) memimpin klasemen dengan marjin yang sangat signifikan, mencatat lebih dari 2.500 kasus pelanggaran. Posisi ini disusul oleh layanan publik (Public), pendidikan (Educational), dan keuangan (Finance).

Mendominasinya sektor kesehatan sebagai target utama peretas memiliki penjelasan ekonomi dan operasional yang sangat logis. Institusi medis secara konstan mengumpulkan, mengelola, dan mentransmisikan Rekam Medis Elektronik (EMR) yang berisi data pribadi tak terubah (seperti riwayat penyakit, jaminan sosial, dan biometrik). Di pasar gelap (dark web), komoditas data kesehatan ini memiliki nilai ekuivalensi finansial yang jauh lebih tinggi dibandingkan data kartu kredit yang dapat diblokir seketika. Selain itu, sifat operasional rumah sakit yang berurusan dengan nyawa manusia menciptakan urgensi absolut; peretas mengetahui bahwa institusi kesehatan akan berada di bawah tekanan ekstrem untuk segera memulihkan sistem mereka ketika terkena serangan ransomware, sehingga probabilitas keberhasilan pemerasan menjadi sangat tinggi.

Di sisi lain, tingginya insiden pada sektor layanan publik dan pendidikan sering kali didorong oleh kombinasi antara volume data berskala raksasa dengan keterbatasan infrastruktur. Institusi pemerintah dan universitas menyimpan jutaan basis data kependudukan dan kemahasiswaan, namun sayangnya sering kali terhambat oleh anggaran keamanan siber yang terbatas (budget constraints) serta penggunaan sistem warisan (legacy systems) yang lambat menerima pembaruan (patching). Celah ini membuka pintu gerbang yang lebar bagi peretas oportunistik. Sementara itu, sektor keuangan dan informasi, meskipun dikenal memiliki lapisan pertahanan siber yang jauh lebih matang, tetap menjadi target persisten karena motivasi ekonomi langsung (pencurian aset likuid) dan spionase korporat.

Dari kacamata arsitektur machine learning, ketimpangan distribusi korban lintas industri ini memberikan wawasan fundamental bagi algoritma. Dalam konstruksi model Random Forest, fitur demografis seperti victim.industry.name bertransformasi menjadi atribut kategorikal yang memiliki daya pemisah (splitting power) yang kuat di dalam simpul-simpul pohon keputusan. Mesin secara otomatis belajar bahwa apabila suatu anomali lalu lintas jaringan terjadi pada entitas di sektor kesehatan atau publik, probabilitas

statistik bahwa insiden tersebut merupakan upaya eksploitasi berlapis yang akan berujung pada Data Breach menjadi jauh lebih tinggi. Dengan demikian, peta distribusi industri ini bukan sekadar statistik deskriptif sebagai pelengkap narasi, melainkan parameter operasional krusial yang membentuk bagaimana algoritma merajut korelasi antara profil korban, nilai data, dan kerentanan sistem secara presisi.




**Grafik 3. Confusion Matrix Model Random Forest**

Hasil evaluasi model secara teknis ditunjukkan melalui *confusion matrix* yang disajikan pada **Grafik 3**, yang merepresentasikan perbandingan antara data insiden aktual dengan hasil prediksi model. Data yang disajikan pada *confusion matrix* ini adalah data hasil olahan (data uji), yang telah melalui tahap *preprocessing* dan SMOTE. Penyajian dalam bentuk matriks memungkinkan pembaca untuk memahami distribusi prediksi benar dan salah secara visual dan sistematis.

Pada kelas *Data Breach*, model berhasil mengklasifikasikan sebanyak 1.289 data secara benar (*True Positives*) dengan hanya 24 data yang salah diprediksi sebagai *No Breach* dan 183 sebagai *Potential Breach*. Tingkat presisi yang sangat tinggi (0,91 atau 91%) dan *recall* sebesar 0,86 pada kelas ini menunjukkan bahwa karakteristik insiden yang berujung pada kebocoran data memiliki pola yang sangat stabil serta mudah dikenali oleh model. Hal ini menandakan bahwa serangan yang sukses memiliki jejak digital (seperti jenis eksploitasi dan aset yang diincar) yang secara konsisten membedakannya dari insiden yang gagal. Secara substantif, temuan ini membuktikan bahwa keberhasilan *data breach* bukanlah fenomena acak, melainkan ditentukan oleh taktik terstruktur. Model bisa menangkap pola tersebut dengan tingkat presisi yang luar biasa.

Pada kelas *Potential Breach*, model mencatat 289 prediksi benar dengan 115 kesalahan klasifikasi yang condong ke kelas *Data Breach*. Secara konseptual, kelas ini merupakan kelas transisional yang mencerminkan fase abu-abu (*grey area*) dalam rantai serangan (*cyber kill-chain*), seperti fase pengintaian (*reconnaissance*) atau pemindaian kerentanan (*scanning*). Kesalahan klasifikasi yang relatif lebih tinggi (dengan presisi 0,61) pada kelas ini bisa dipahami sebagai konsekuensi dari ambiguitas perilaku lalu lintas jaringan. Insiden pada fase ini umumnya telah menunjukkan indikator kompromi (*Indicator of Compromise*), namun masih sulit dipastikan apakah akan berujung pada ekstraksi data atau hanya sekadar intrusi pasif. Tumpang tindih karakteristik dengan kelas *Data Breach* menjadi tidak terhindarkan. Temuan ini justru memperkuat argumen bahwa serangan siber adalah proses multi-tahap yang berkelanjutan, bukan sekadar peristiwa biner.

Sementara itu, pada kelas *No Breach*, model berhasil mengklasifikasikan sebanyak 145 data secara benar dengan *recall* mencapai 0,94 (94%). Dominasi prediksi benar dari sisi sensitivitas pada kelas ini membuktikan bahwa model memiliki kemampuan yang sangat kuat dalam mengenali pola serangan yang berhasil dipatahkan oleh sistem pertahanan

(seperti *firewall* atau *Intrusion Prevention System*). Tingginya *recall* pada kelas ini mengindikasikan bahwa anomali serangan yang gagal menghasilkan pola lalu lintas yang spesifik dan mudah diidentifikasi sebelum mencapai aset inti.

Secara keseluruhan, model mencapai tingkat akurasi ( *Accuracy* ) sebesar 83,12% (0,8312). Nilai akurasi yang tinggi ini didukung oleh keseimbangan performa makro dan berbobot (*weighted avg F1-score* 0,83). Hal ini menunjukkan bahwa mayoritas data berhasil diklasifikasikan dengan tepat, sementara nilai yang seimbang menunjukkan bahwa model mampu menjaga ekuilibrium antara *precision* dan *recall*. Hal ini sangat penting dalam konteks data keamanan siber yang secara inheren tidak seimbang. Kombinasi metrik tersebut menegaskan bahwa performa model tidak bias terhadap kelas mayoritas (hanya menebak *Data Breach*), melainkan benar-benar mempelajari fitur. Dengan demikian, hasil evaluasi ini menunjukkan bahwa *Random Forest* merupakan model yang reliabel untuk membedah perilaku serangan siber.

**Analisis Stabilitas dan Reliabilitas Model**

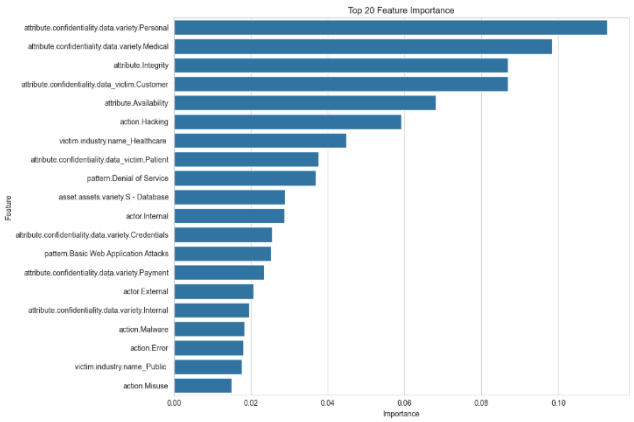
Selain kinerja klasifikasi dalam memisahkan kelas ancaman, stabilitas dan reliabilitas model menjadi aspek fundamental dalam menilai kualitas serta kelayakan operasional dari hasil penelitian prediktif keamanan siber. Algoritma Random Forest memiliki keunggulan komparatif yang terletak pada mekanisme ensemble learning. Pada arsitektur ini, keputusan akhir tidak bergantung pada satu alur logika saja, melainkan didapatkan dari agregasi hasil prediksi ratusan pohon keputusan (dalam penelitian ini dioptimalkan menggunakan parameter n_estimators=500) yang dibangun secara acak. Pendekatan ini mengimplementasikan teknik Bootstrap Aggregating (Bagging) dan pemilihan fitur acak (random feature selection) pada setiap simpul (node). Mekanisme ini secara spesifik dirancang untuk mereduksi secara drastis risiko overfitting yang sangat sering terjadi pada model klasifikasi tunggal ketika dipaksa menghadapi log jaringan berskala raksasa. Berbeda dengan algoritma Decision Tree konvensional yang sangat sensitif terhadap fluktuasi data latih dan cenderung menghafal data, Random Forest membagi varians sehingga menghasilkan prediksi yang jauh lebih stabil. Dalam konteks data intelijen ancaman yang kompleks dan multidimensional, stabilitas ansambel ini menjadi indikator krusial bahwa tingginya metrik evaluasi yang dicapai tidak bersifat kebetulan (noise), melainkan benar-benar merepresentasikan logika serangan siber yang terstruktur.

Hasil evaluasi mengonfirmasi bahwa model mampu mempertahankan performa yang konsisten melintasi spektrum ancaman yang luas. Kemampuan generalisasi (generalization capability) ini menjadi metrik yang tak kalah penting dari sekadar angka akurasi murni. Konsistensi ini tercermin dari distribusi akurasi yang tetap terjaga meskipun data uji (test set) mencakup berbagai variasi vektor serangan tingkat lanjut—mulai dari eksploitasi kerentanan aplikasi (Hacking), injeksi Malware, hingga manipulasi psikologis (Social Engineering)—yang berasal dari latar belakang industri korban yang sangat beragam. Dataset VCDB, sebagai repositori global, merekam insiden tanpa memandang batasan yurisdiksi, sehingga mencerminkan perbedaan infrastruktur arsitektur jaringan, skala organisasi (mulai dari bisnis menengah hingga kelas Enterprise), dan maturitas keamanan (security posture) yang sangat bervariasi. Model klasifikasi yang tidak reliabel cenderung akan mengalami penurunan performa secara drastis (performance degradation) atau gagal menangkap pola umum ketika dihadapkan pada serangan siber varian baru (zero-day) maupun taktik hibrida. Namun, hasil pengujian membuktikan bahwa Random Forest mampu mengakomodasi keragaman (heterogeneity) dan dimensi fitur yang tinggi tersebut tanpa sedikit pun mengorbankan fungsi presisi deteksi dini.

Pada tingkatan yang lebih strategis, stabilitas model merupakan prasyarat mutlak sebelum melangkah pada analisis feature importance (tingkat kepentingan fitur). Validitas dari identifikasi variabel operasional mana yang memicu kebocoran data sangat bergantung pada kekokohan arsitektur prediktif di belakangnya. Interpretasi mengenai variabel yang menjadi titik lemah sistem hanya bernilai valid secara ilmiah apabila model mempunyai kinerja yang konsisten dan tidak fluktuatif ketika divalidasi silang dengan porsi data yang berbeda. Apabila model tidak stabil, maka nilai kepentingan variabel berpotensi sekadar mencerminkan anomali statistik matematis atau noise sesaat, bukan memetakan pola substantif dari taktik peretasan di dunia nyata. Model yang labil akan menghasilkan urutan prioritas fitur yang berubah-ubah secara ekstrem, yang pada akhirnya justru akan menyesatkan arah investasi mitigasi bagi pengambil keputusan. Oleh karena itu, tingkat stabilitas yang berhasil dicapai dalam penelitian ini—yakni 83,12% akurasi pada data validasi eksternal, dikombinasikan dengan metrik recall yang superior—menunjukkan legitimasi metodologis yang sangat kokoh. Tingkat keandalan ini memastikan bahwa analisis faktor utama yang memengaruhi probabilitas data breach (seperti terekspos nya data medis dan personal) adalah temuan empiris yang sepenuhnya dapat dipertanggungjawabkan untuk merancang postur pertahanan siber di masa depan.

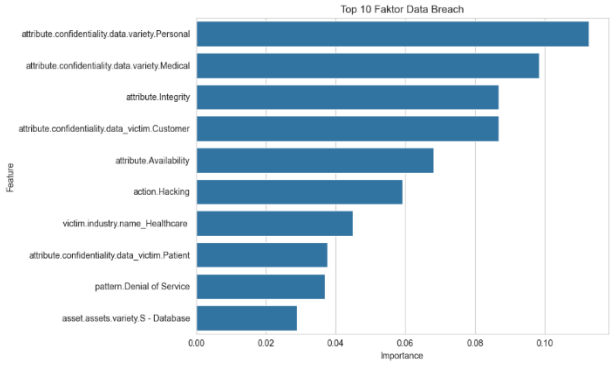
**Analisis Faktor Utama Berdasarkan Feature Importance**

Analisis *feature importance* bertujuan untuk mengidentifikasi variabel-variabel operasional, teknis, dan demografis yang paling berkontribusi terhadap keberhasilan suatu eksploitasi data. Data yang diekstrak mencerminkan kontribusi relatif setiap variabel dalam membelah (*splitting*) node pada pohon keputusan, merepresentasikan seberapa besar peran variabel tersebut dalam membedakan apakah suatu sistem akan jebol atau bertahan. Analisis ini bertindak sebagai jembatan langsung antara evaluasi teknis dan pemahaman risiko strategis.




**Grafik 4. Top 20 Feature Importance dalam Prediksi Data Breach**

Berdasarkan **Grafik 4** dan secara lebih spesifik **Grafik 5**, variabel attribute.confidentiality.data.variety.Personal memiliki tingkat kepentingan absolut tertinggi dengan skor melampaui 0,11. Temuan ini menyingkap fakta bahwa keberadaan data pribadi (*Personally Identifiable Information*) merupakan magnit utama dan prediktor terkuat terjadinya *data breach*. Secara ekonomis dan psikologis bagi peretas, data personal adalah komoditas dengan likuiditas tertinggi di forum kejahatan siber (*dark web*). Keterlibatan data personal berfungsi sebagai katalisator yang mengubah intrusi jaringan biasa menjadi insiden pencurian data berskala penuh.

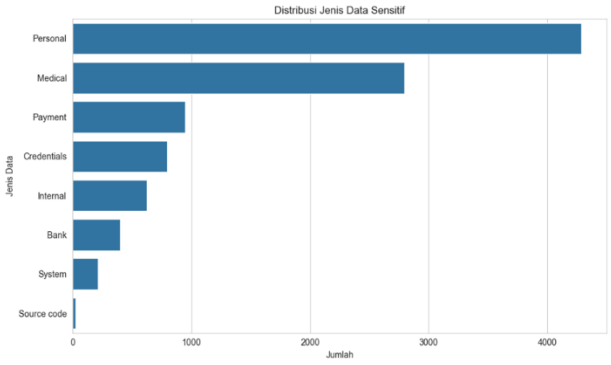



**Grafik 5. Sepuluh Faktor Dominan Penyebab Data Breach**

Variabel attribute.confidentiality.data.variety.Medical menempati posisi kedua dengan skor mendekati 0,10. Hal ini berkorelasi langsung dengan tingginya serangan pada industri *Healthcare*. Nilai rekam medis sangat tinggi karena sifatnya yang abadi (imutabilitas biologis); pasien tidak bisa mengganti golongan darah atau riwayat penyakit mereka semudah mengganti kata sandi. Fakta ini menegaskan bahwa tingkat risiko sebuah organisasi sangat didikte oleh nilai intrinsik dari data yang mereka kelola.

Selain itu, variabel attribute.Integrity dan attribute.Availability masuk dalam jajaran lima besar. Hal ini membuktikan bahwa serangan siber modern tidak hanya bertujuan mencuri data (kerahasiaan), tetapi sering kali melibatkan manipulasi sistem (seperti instalasi *backdoor* yang merusak integritas) atau melumpuhkan akses layanan (seperti *ransomware* atau DDoS yang menghancurkan ketersediaan). Kombinasi variabel ini mencerminkan taktik pemerasan ganda (*double extortion*) yang marak dilakukan sindikat peretas.

Dari sisi metode, variabel action.Hacking dan pattern.Denial of Service menjadi vektor serangan paling dominan. Ini menunjukkan bahwa meskipun rekayasa sosial (*phishing*) sering dibicarakan, eksploitasi kelemahan teknis pada perangkat keras maupun lunak (seperti kerentanan *server* atau aplikasi web dasar) tetap menjadi jalur penetrasi yang paling menentukan keberhasilan kebocoran data.




**Grafik 6. Distribusi Spesifik Jenis Data Sensitif**

Eksplorasi lebih mendalam terhadap aset yang terkompromi pada **Grafik 5** memperlihatkan secara gamblang bahwa akumulasi data *Personal* (mencapai lebih dari 4.000 kasus) dan *Medical* (mendekati 3.000 kasus) mendominasi lanskap kebocoran, jauh meninggalkan data *Payment* (keuangan), *Credentials* (kata sandi), dan *Internal*. Secara keseluruhan, hasil analisis *feature importance* dan distribusi data sensitif ini membuktikan bahwa keberhasilan *data breach* adalah sebuah konvergensi mematikan antara aset bernilai tinggi (data personal/medis) dengan vektor teknis yang rapuh (integritas basis data dan taktik peretasan). Tidak ada satu anomali tunggal yang menyebabkan kebocoran; ini adalah fenomena rantai kegagalan yang multidimensional.

**Pembahasan**

**Kaitan Hasil Penelitian dengan Konsep Keamanan Informasi**

Hasil penelitian ini menunjukkan keterkaitan yang sangat kuat antara temuan empiris *machine learning* dengan konsep fundamental keamanan siber, khususnya Triad CIA (*Confidentiality, Integrity, Availability*). Dominasi variabel attribute.Integrity, attribute.Availability, dan attribute.confidentiality.data.variety.Personal dalam analisis fitur secara langsung mendukung aksioma utama bahwa perlindungan informasi tidak bisa hanya berfokus pada satu pilar. Temuan ini memperlihatkan bahwa keberhasilan peretas untuk mengekstrak data (melanggar kerahasiaan) hampir selalu didahului atau dibarengi dengan pelanggaran terhadap integritas sistem (misalnya, meretas basis data untuk mengubah hak akses) dan ketersediaan layanan. Dengan kata lain, hasil penelitian ini mengonfirmasi bahwa kerangka analitik *cyber kill-chain* beroperasi secara nyata dalam dataset VCDB, di mana eksploitasi berjalan secara sekuensial.

Peran menonjol dari industri kesehatan (victim.industry.name_Healthcare) dan target individu (attribute.confidentiality.data_victim.Patient) memperluas pemahaman kita terhadap konsep *Risk-Based Security* (Keamanan Berbasis Risiko). Hasil penelitian membuktikan bahwa tidak semua data diciptakan setara di mata *threat actors*. Institusi yang memegang data medis dan personal mengalami tekanan serangan (*attack surface*) yang asimetris dibandingkan industri manufaktur atau hiburan. Temuan ini sejalan dengan asumsi ekonomi kejahatan siber yang menyatakan bahwa peretas adalah entitas rasional yang mengkalkulasi *Return on Investment* (ROI) dari operasi mereka. Semakin berharga data yang disimpan, semakin canggih dan persisten metode *Hacking* yang akan dikerahkan. Hal ini membuktikan bahwa strategi mitigasi organisasi tidak boleh lagi bersifat "satu ukuran untuk semua" (*one-size-fits-all*), melainkan harus difokuskan pada perlindungan isolatif terhadap aset-aset berharga tinggi tersebut.

Selain itu, kemunculan variabel asset.assets.variety.S - Database sebagai salah satu target aset tertinggi menunjukkan bahwa pusat repositori data (*core databases*) masih menjadi piala suci (*holy grail*) bagi penyerang. Berbeda dengan dekade lalu di mana peretas sering kali hanya merusak tampilan situs web (*defacement*), kini motivasi mereka murni berorientasi pada akuisisi basis data massal. Temuan ini menggarisbawahi kelemahan arsitektur perimeter tradisional yang sering kali gagal melindungi lapisan inti basis data setelah jaringan luar berhasil ditembus.

**Perbandingan dengan Penelitian Terdahulu**

Hasil penelitian ini menunjukkan tingkat kesesuaian dan eskalasi temuan yang kuat jika dikalibrasi dengan literatur terdahulu. Signifikansi variabel action.Hacking dan pattern.Denial of Service sangat sejalan dengan penelitian (Mareedu, 2023) yang mengemukakan bahwa ancaman teknis dan intrusi sistem tingkat lanjut masih menjadi alat utama eksploitasi modern yang memerlukan integrasi *Threat Intelligence*. Konsistensi ini mengindikasikan bahwa serangan berbasis injeksi teknis dan eksploitasi perangkat lunak

tidak pernah kehilangan relevansinya, dan model *machine learning* berhasil merekam anomali ini sebagai prediktor kuat.

Namun, penelitian ini juga memperluas batas pemahaman yang pernah diajukan oleh (Folino et al., 2023) mengenai deteksi anomali perilaku (*User Behavior Anomaly*). Meskipun penelitian Folino menitikberatkan pada perilaku *insider* dan anomali pengguna, hasil *feature importance* dalam penelitian ini menunjukkan bahwa variabel actor.Internal dan action.Misuse memang memiliki andil, namun urgensinya masih berada di bawah taktik peretasan eksternal (actor.External dan action.Hacking) serta tipe data yang dicuri. Hal ini mengonfirmasi bahwa meskipun ancaman dari dalam (*insider threat*) itu nyata dan merusak, volume dan probabilitas insiden yang berujung pada *data breach* berskala besar secara global tetap didominasi oleh penetrasi pihak luar (*external threat actors*).

Penelitian ini juga sejalan sekaligus mempertajam temuan (Hu et al., 2023) yang menganalisis risiko rantai pasokan. Karena banyak basis data modern dan sistem informasi dikelola melalui komputasi awan atau vendor perangkat lunak pihak ketiga, kerentanan yang terdeteksi pada *asset database* mengisyaratkan tingginya risiko dari kegagalan pihak ketiga dalam mengamankan integrasi data. Kesesuaian temuan ini menunjukkan bahwa model prediktif berbasis *Random Forest* memiliki fleksibilitas untuk memvalidasi hipotesis dari berbagai domain studi keamanan, baik itu studi berfokus teknis, kebijakan tata kelola, maupun manajemen risiko pihak ketiga.

**Implikasi Teoretis dan Praktis**

Hasil penelitian ini memberikan implikasi teoretis yang signifikan terhadap pengembangan model intelijen ancaman siber (*Cyber Threat Intelligence*). Secara teoritis, penelitian ini membuktikan bahwa pemodelan prediktif untuk *data breach* tidak bisa hanya bergantung pada log jaringan atau identifikasi IP anomali semata (pendekatan *network-centric*), melainkan harus secara definitif mengintegrasikan dimensi atribut data (*data-centric approach*). Dominasi nilai prediktif dari fitur jenis data (*Personal* dan *Medical*) memaksa literatur deteksi intrusi untuk berevolusi; mesin kecerdasan buatan harus diajarkan untuk memahami **"apa"** yang sedang diakses oleh penyerang dengan bobot yang sama pentingnya dengan **"bagaimana"** cara mereka mengaksesnya. Model ini memperkaya diskursus akademik mengenai pergeseran paradigma keamanan dari perlindungan infrastruktur menuju perlindungan aset informasi.

Dari sisi praktis, hasil penelitian ini memberikan panduan strategis yang sangat tajam bagi *Chief Information Security Officers* (CISO), arsitek keamanan, dan pengambil kebijakan IT. Pertama, dengan terbuktinya kerentanan mutlak pada kredensial, hak akses, dan peretasan basis data, organisasi wajib meninggalkan model keamanan perimeter klasik (seperti perlindungan *firewall* berbasis batas jaringan) dan beralih sepenuhnya pada arsitektur **Zero Trust**. Model ini berasumsi bahwa jaringan telah dikompromikan, sehingga verifikasi berkelanjutan menggunakan *Multi-Factor Authentication* (MFA) diwajibkan setiap kali terjadi upaya akses menuju pangkalan data sensitif, dari mana pun asalnya.

Kedua, besarnya kontribusi taktik eksploitasi multi-tahap mengamanatkan implementasi otomatisasi keamanan melalui *Security Information and Event Management* (SIEM) generasi terbaru yang diperkuat algoritma ML (seperti yang dieksperimenkan dalam studi ini). SIEM ini tidak hanya mengumpulkan log, tetapi harus mampu melakukan korelasi proaktif antara upaya *Hacking* awal dengan anomali pembacaan data *Personal*. Waktu tanggap insiden (*Incident Response Time*) harus ditekan sebelum peretas mencapai fase eksfiltrasi (pengunduhan data ke luar jaringan).

Ketiga, bagi industri dengan regulasi ketat (khususnya *Healthcare* dan *Finance*), temuan ini adalah alarm regulatori. Mengingat data medis dan personal menjadi incaran

utama, strategi mitigasi teknis level akhir seperti enkripsi *data-at-rest* standar militer (misalnya AES-256), *Dynamic Data Masking*, serta tokenisasi harus dijadikan mandat operasional, bukan sekadar pelengkap kepatuhan audit. Dengan enkripsi menyeluruh pada penyimpanan basis data, sekalipun lapisan integritas (attribute.Integrity) runtuh dan peretas berhasil mencuri basis data, informasi medis dan pribadi tersebut akan tetap berupa *ciphertext* acak yang nir-nilai di pasar gelap, secara efektif menetralisir tujuan utama dari kejahatan siber tersebut.

## KESIMPULAN

Penelitian ini berhasil mengidentifikasi faktor-faktor risiko utama yang memengaruhi terjadinya data breach serta membangun model prediksi menggunakan algoritma Random Forest berbasis Verizon Community Database (VCDB). Hasil pengujian menunjukkan bahwa model mampu mengklasifikasikan insiden keamanan siber ke dalam kategori Data Breach, Potential Breach, dan No Breach dengan akurasi sebesar 83,12%. Penerapan teknik SMOTE juga terbukti efektif dalam mengatasi ketidakseimbangan kelas sehingga model dapat mengenali pola ancaman secara lebih baik dan menghasilkan performa klasifikasi yang stabil.

Analisis feature importance menunjukkan bahwa data personal dan data medis merupakan faktor yang paling berpengaruh terhadap terjadinya data breach, diikuti oleh aspek integritas dan ketersediaan sistem serta aktivitas hacking sebagai vektor serangan utama. Selain itu, sektor Healthcare, Public, dan Educational menjadi industri yang paling rentan terhadap insiden keamanan siber. Temuan ini menunjukkan bahwa data breach merupakan hasil interaksi berbagai faktor teknis dan operasional, sehingga organisasi perlu menerapkan pendekatan keamanan yang proaktif, berbasis risiko, dan didukung oleh teknologi machine learning untuk meningkatkan kemampuan deteksi serta mitigasi ancaman siber.

## DAFTAR PUSTAKA


Ahsan, M., Nygard, K. E., Gomes, R., Chowdhury, M., & Rifat, N. (2022). Cybersecurity Threats and Their Mitigation Approaches Using Machine Learning — A Review. 527–555.

Alkhudaydi, O. A., Krichen, M., & Alghamdi, A. D. (2023). A Deep Learning Methodology for Predicting Cybersecurity Attacks on the Internet of Things.

Almahmoud, Z., Yoo, P. D., Alhussein, O., Farhat, I., & Damiani, E. (2023). OPEN A holistic and proactive approach to forecasting cyber threats. Scientific Reports, 1–15. https://doi.org/10.1038/s41598-023-35198-1

Dalal, S., Lilhore, U. K., Faujdar, N., Simaiya, S., Ayadi, M., & Almujally, N. A. (2023). Next - generation cyber attack prediction for IoT systems : leveraging multi - class SVM and optimized CHAID decision tree. https://doi.org/10.1186/s13677-023-00517-4

Dalal, S., Manoharan, P., Kumar, L. U., Seth, B., Mohammed, D., Simaiya, S., Hamdi, M., & Raahemifar, K. (2023). Extremely boosted neural network for more accurate multi - stage Cyber attack prediction in cloud computing environment. https://doi.org/10.1186/s13677-022-00356-9

Elluri, L., Mandalapu, V., Vyas, P., & Roy, N. (2023). REVIEW ARTICLE Recent Advancements in Machine Learning For Cybercrime Prediction. 2020, 1–21.

Folino, G., Godano, C. O., & Pisani, F. S. (2023). detection and classification for cybersecurity. The Journal of Supercomputing, 79(11), 11660–11683. https://doi.org/10.1007/s11227-023-05049-x

Hammad, A. A. (2024). Random Forest and LSTM Hybrid Model for Detecting DDoS Attacks in Healthcare IoT Networks. 1(2), 1–8.

Hu, K., Levi, R., Yahalom, R., & Zerhouni, E. G. (2023). Supply Chain Characteristics as Predictors

of Cyber Risk : A Machine-Learning Assessment. 1–9.
Li, J., Xiao, W., & Zhang, C. (2023). Data security crisis in universities: identification of key factors affecting data breach incidents. 2023. https://doi.org/10.1057/s41599-023-01757-0
Mareedu, A. (2023). Data-Driven Cybersecurity : ML-Based Threat Intelligence and Prediction Systems. 1(3), 212–223. https://doi.org/10.56472/25838628/IJACT-V1I3P122
Markovic, T., Leon, M., Buffoni, D., & Punnekkat, S. (2024). Random forest with differential privacy in federated learning framework for network attack detection and classification. 8132–8153. https://doi.org/10.1007/s10489-024-05589-6
Roy, P., Tech, V., Lanus, E., Tech, V., Freeman, L., & Tech, V. (2023). A Survey of Data Security : Practices from Cybersecurity and Challenges of Machine Learning arXiv : 2310 . 04513v3 [ cs . CR ] 4 Dec 2023.
Sarker, I. H. (2022). Machine Learning for Intelligent Data Analysis and Automation in Cybersecurity : Current and Future Prospects. Annals of Data Science, 10(6), 1473–1498. https://doi.org/10.1007/s40745-022-00444-2
Saxena, D., Gupta, I., Gupta, R., Singh, A., & Wen, X. (2023). An AI-Driven VM Threat Prediction Model for. 1–13.