Hello World
Jurnal Ilmu Komputer
https://jurnal.ilmubersama.com/index.php/hello_world

Artikel Penelitian

Sistem Deteksi Otomatis Berbasis Machine Learning untuk Jejak Digital dalam Analisis Forensik Siber

Pius Deski Manalu [1], Ammar Yasir Nasution [2] , Ekoliyono Wahyu Sumantri [3], Mohd Azlishah Othman [4], Muhammad Nasri Gea[5]*

1) *STIE Professional Manajemen College Indonesia*
2) *Al-Azhar University.*
3) *Universitas Potensi Utama*
4) *Department of Engineering Technology, Universiti Teknikal Malaysia Melaka*
5) *Akademi Manajemen Informatika Dan Komputer ITMI*

INFORMASI ARTIKEL

Diterima Redaksi: -
Revisi Akhir: -
Diterbitkan Online: -

A B S T R A K

Meningkatnya volume dan heterogenitas bukti digital mendorong kajian deteksi jejak digital berbasis machine learning. Penelitian ini mengkaji integrasi fitur dari log sistem, lalu lintas jaringan, dan artefak sistem pada data simulasi dengan Random Forest (RF) dan Support Vector Machine (SVM). Naskah melaporkan 15.000 sampel dan pembagian 80:20; RF memperoleh akurasi 96,2%, presisi 95,8%, dan recall 94,7%, sedangkan SVM memperoleh akurasi 92,3%. Uji McNemar yang dilaporkan (χ^2 = 14,73; $p < 0,001$) hanya menyangkut perbandingan RF dan SVM, bukan pembanding analis manual. Studi ablasi menunjukkan selisih akurasi 7,8 poin persentase antara model multisumber dan varian log saja dalam evaluasi yang dilaporkan. Namun, distribusi kelas, definisi unit sampel lintas sumber, pemisahan kelompok yang mencegah kebocoran data, rincian lingkungan simulasi, serta ketidakpastian hasil belum terdokumentasi memadai. Karena itu, angka kinerja perlu dipandang sebagai hasil awal pada simulasi, bukan bukti kesiapan penggunaan forensik operasional. Pengujian ulang berbasis sesi/host/waktu, pelaporan AUC-ROC dan PR-AUC, pembanding tambahan, serta validasi pada dataset eksternal diperlukan sebelum menarik kesimpulan generalisasi.

KATA KUNCI

machine learning; forensik siber; deteksi jejak digital; random forest; investigasi otomatis

KORESPONDENSI

Ammar Yasir Nasution

E-mail: yasirnasti0396@gmail.com

https://doi.org/10.12345/xxxxx

Attribution-ShareAlike 4.0 International Some rights reserved

Hello World
Jurnal Ilmu Komputer
https://jurnal.ilmubersama.com/index.php/hello_world

1. PENDAHULUAN

Kemajuan pesat teknologi informasi dan komunikasi telah memperluas ruang lingkup aktivitas siber, baik untuk kegiatan yang sah maupun kejahatan. Dalam beberapa tahun terakhir, kejahatan siber semakin kompleks, mulai dari serangan ransomware dan kebocoran data hingga manipulasi sistem serta penyebaran jejak digital pada lingkungan yang heterogen. Integrasi data digital berskala besar dan perangkat yang saling terhubung menimbulkan tantangan besar bagi proses investigasi forensik, termasuk pengumpulan artefak digital, korelasi jejak, dan rekonstruksi peristiwa [1]. Pendekatan manual dan semiotomatis menghadapi tantangan skalabilitas pada data berukuran besar dan heterogen [2]; besaran tingkat kesalahan analis perlu ditetapkan melalui studi empiris yang melaporkan rancangan pengukurannya. Keterbatasan ini menegaskan perlunya sistem forensik otomatis dan cerdas yang mampu menangani kecepatan, volume, serta keragaman bukti digital modern.

Forensik digital telah berkembang menjadi disiplin sistematis yang bertujuan mengidentifikasi, memulihkan, menganalisis, dan mendokumentasikan bukti digital yang dapat dipertanggungjawabkan secara hukum berdasarkan kerangka prosedur standar [2]. Namun, pendekatan tradisional sulit mengikuti skala data digital modern. Literatur terkini menunjukkan bahwa kecerdasan buatan, khususnya machine learning, menawarkan solusi untuk mempercepat, memperluas, dan meningkatkan akurasi proses forensik digital [3]. Teknik machine learning memungkinkan ekstraksi fitur otomatis, klasifikasi aktivitas mencurigakan, dan rekonstruksi kronologi peristiwa siber [4]. Tinjauan sistematis penerapan machine learning dalam forensik siber, seperti yang dilakukan Ali et al. (2022), mengonfirmasi semakin kuatnya bukti yang mendukung pendekatan berbasis machine learning pada berbagai ranah forensik, termasuk analisis malware, deteksi intrusi jaringan, dan forensik citra [12]. Walaupun demikian, bidang ini masih menghadapi tantangan implementasi penting yang membatasi penerapannya dalam praktik.

Kesenjangan penelitian saat ini mencakup beberapa aspek. Pertama, peningkatan volume data digital yang sangat besar membuat metode manual tidak efisien, sedangkan sebagian besar penelitian machine learning forensik masih terbatas pada citra atau analisis malware, bukan deteksi jejak aktivitas pengguna secara menyeluruh dan lintas sumber [5]. Kedua, keterbatasan dataset forensik yang terbuka, berlabel, dan relevan secara praktis menghambat pengembangan model yang siap diterapkan serta menimbulkan keraguan terhadap generalisasi hasil dari lingkungan simulasi ke kasus forensik nyata [6]. Ketiga, terdapat kesenjangan antara penelitian akademik dan praktik lapangan terkait implementasi, akuntabilitas, dan keabsahan hukum bukti digital, termasuk keterbatasan transparansi, interpretabilitas, serta kesiapan prosedural model [7]. Keempat, meskipun Ali et al. (2022) telah memberikan tinjauan sistematis tentang machine learning dalam forensik siber, hasilnya menunjukkan bahwa integrasi serentak berbagai sumber data forensik, khususnya log sistem, lalu lintas jaringan, dan artefak sistem, dalam satu kerangka deteksi masih sangat sedikit diteliti [12].

Penelitian ini secara langsung mengatasi kesenjangan tersebut dengan mengembangkan dan mengevaluasi model machine learning untuk deteksi otomatis jejak digital melalui integrasi fitur lintas sumber. Perbedaan utama dari kerangka yang ditinjau Ali et al. (2022) terletak pada tiga aspek. Pertama, penelitian ini tidak hanya menerapkan machine learning pada satu ranah data forensik, tetapi mengintegrasikan tiga sumber heterogen secara bersamaan, yaitu log sistem, rekaman lalu lintas jaringan, dan artefak sistem, sehingga pengenalan pola perilaku menjadi lebih kaya dan komprehensif. Kedua, penelitian ini menggunakan baseline prosedur forensik digital standar sebagaimana didefinisikan Montasari et al. (2022), sebagai kerangka konseptual alur investigasi, bukan sebagai baseline numerik yang telah diuji. Ketiga, dengan tetap mengakui keterbatasan dataset simulasi, penelitian ini menyediakan kerangka dasar untuk divalidasi pada lingkungan forensik nyata. Arti penting penelitian ini mencakup potensi percepatan pemrosesan dan dukungan bagi penyaringan bukti digital, yang masih memerlukan evaluasi operasional dan pemeriksaan manusia. Random Forest dan Support Vector Machine dipadukan dengan ekstraksi fitur lintas sumber untuk membentuk kerangka deteksi otomatis yang dapat diterapkan pada platform forensik nyata.

2. METODOLOGI

2.1. Desain dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen terapan untuk mengevaluasi efektivitas algoritma machine learning dalam mendeteksi jejak digital secara otomatis. Teknik supervised learning, yaitu Random Forest dan Support Vector Machine, digunakan untuk mengidentifikasi dan mengklasifikasikan pola aktivitas digital yang berkaitan dengan insiden siber. Desain eksperimen mengikuti alur sistematis yang meliputi pengumpulan data, prapemrosesan, pelatihan model, pengujian, dan evaluasi kinerja.

Prosedur forensik manual yang dibahas dalam literatur [2] mencakup identifikasi, pengumpulan, pemeriksaan, dan analisis bukti. Prosedur tersebut digunakan dalam naskah ini sebagai rujukan konseptual, bukan sebagai kelompok pembanding eksperimental yang setara dengan model. Dokumen penelitian yang tersedia belum menjelaskan jumlah

https://doi.org/10.12345/xxxxx Attribution-ShareAlike 4.0 International Some rights reserved

analis, pengalaman, tugas, alokasi sampel, protokol pencatatan waktu, ataupun hasil anotasi manual per sampel. Karena itu, angka metrik dan waktu untuk metode manual yang sebelumnya dicantumkan tidak dipertahankan sebagai hasil eksperimen. Angka tingkat kesalahan 12–18% juga tidak digunakan karena verifikasi terhadap sumber dan konteks pengukurannya belum tersedia.

2.2. Pengumpulan Data dan Persiapan Dataset

Jejak digital dikumpulkan dari lingkungan siber simulasi yang merepresentasikan aktivitas normal dan mencurigakan. Lingkungan simulasi dirancang berdasarkan taksonomi serangan nyata, meliputi akses tanpa izin, pergerakan lateral, eksfiltrasi data, dan penyebaran ransomware, untuk menjaga validitas ekologis dataset. Penggunaan data simulasi diakui sebagai keterbatasan karena mungkin belum sepenuhnya menggambarkan kecanggihan teknik pengelakan di dunia nyata. Namun, pendekatan ini mengikuti praktik yang telah digunakan dalam penelitian machine learning forensik [5][6] dan menyediakan dasar yang terkendali serta dapat direplikasi untuk mengevaluasi algoritma. Validasi lanjutan pada lingkungan operasional nyata direncanakan sebagai kelanjutan penelitian. Deskripsi yang tersedia belum merinci jumlah host, sistem operasi target, jumlah sesi, durasi pengumpulan, perangkat pembangkitan serangan, maupun prosedur menghasilkan aktivitas normal. Komponen registry menunjuk pada artefak target yang mendukung registry dan tidak boleh disamakan dengan Ubuntu 22.04 yang disebut sebagai sistem operasi workstation pelatihan. Informasi lingkungan tersebut harus dilengkapi dari catatan eksperimen sebelum replikasi dapat dinilai.

Dataset terdiri atas 15.000 sampel dari tiga sumber utama. Sumber pertama adalah berkas log sistem yang memuat catatan autentikasi pengguna, jejak eksekusi proses, dan perubahan registry. Sumber kedua adalah rekaman lalu lintas jaringan menggunakan Wireshark untuk menganalisis header paket, metadata koneksi, dan pola komunikasi khusus protokol. Sumber ketiga adalah artefak sistem yang diperoleh dengan perangkat forensik seperti FTK Imager dan Autopsy, meliputi riwayat peramban, berkas sementara, dan pola penggunaan aplikasi. Integrasi tiga sumber ini dirancang untuk mengatasi keterbatasan pendekatan sumber tunggal, seperti model berbasis jaringan saja oleh Kara dan Varol (2022), melalui korelasi perilaku lintas sumber yang menghasilkan profil bukti forensik lebih lengkap. Unit analisis untuk satu "sampel" belum didefinisikan secara operasional dalam bahan yang tersedia: apakah satu host, satu sesi, satu peristiwa, atau satu jendela waktu. Aturan penyelarasan cap waktu dan identitas sumber log, pcap, serta artefak perlu disertakan; tanpa aturan itu, klaim korelasi lintas sumber masih bersifat terbatas.

Prapemrosesan data meliputi beberapa tahap penting. Tahap pertama adalah pembersihan data untuk menghapus rekaman yang tidak lengkap dan menjaga integritas data. Tahap kedua adalah normalisasi menggunakan Min-Max Scaling untuk menyeragamkan skala fitur pada jenis data yang heterogen. Tahap ketiga adalah pemberian label untuk membedakan aktivitas normal dan potensi insiden keamanan, yang diverifikasi dengan catatan ground truth serangan simulasi. Untuk mencegah data leakage pada evaluasi berikutnya, pemilihan fitur, penskalaan, dan PCA harus di-fit hanya pada subset pelatihan di setiap lipatan validasi silang, lalu diaplikasikan tanpa fit ulang pada subset validasi/uji. Dokumen yang tersedia belum membuktikan bahwa urutan ini telah dilaksanakan pada eksperimen yang dilaporkan.

Ekstraksi fitur menggunakan beberapa teknik untuk mengidentifikasi indikator forensik paling penting dari jejak digital mentah. Information Gain digunakan untuk mengurutkan fitur berdasarkan daya diskriminasinya terhadap kelas target dan memilih atribut paling relevan dari setiap sumber data. Uji Chi-Square digunakan untuk menilai independensi statistik fitur kategorikal, terutama pada log sistem dan artefak peramban. Untuk fitur lalu lintas jaringan berdimensi tinggi, Principal Component Analysis (PCA) digunakan guna mengurangi dimensi sambil mempertahankan 95% varians total, sehingga beban komputasi berkurang tanpa menghilangkan informasi prediktif. Gabungan fitur berfokus pada frekuensi akses log, pola koneksi jaringan, urutan waktu, dan cap waktu perubahan berkas sebagai indikator utama dari ketiga sumber data.

Untuk memvalidasi keunggulan integrasi multisumber secara kuantitatif, dilakukan eksperimen ablasi terkendali. Tiga varian model dievaluasi, yaitu: (1) model berbasis log saja, (2) model berbasis lalu lintas jaringan saja, dan (3) model berbasis artefak sistem saja. Setiap varian dibandingkan dengan model lengkap yang mengintegrasikan ketiga sumber. Hasil analisis ablasi disajikan pada Tabel 2 dan menunjukkan peningkatan akurasi terukur yang berasal dari integrasi fitur lintas sumber.

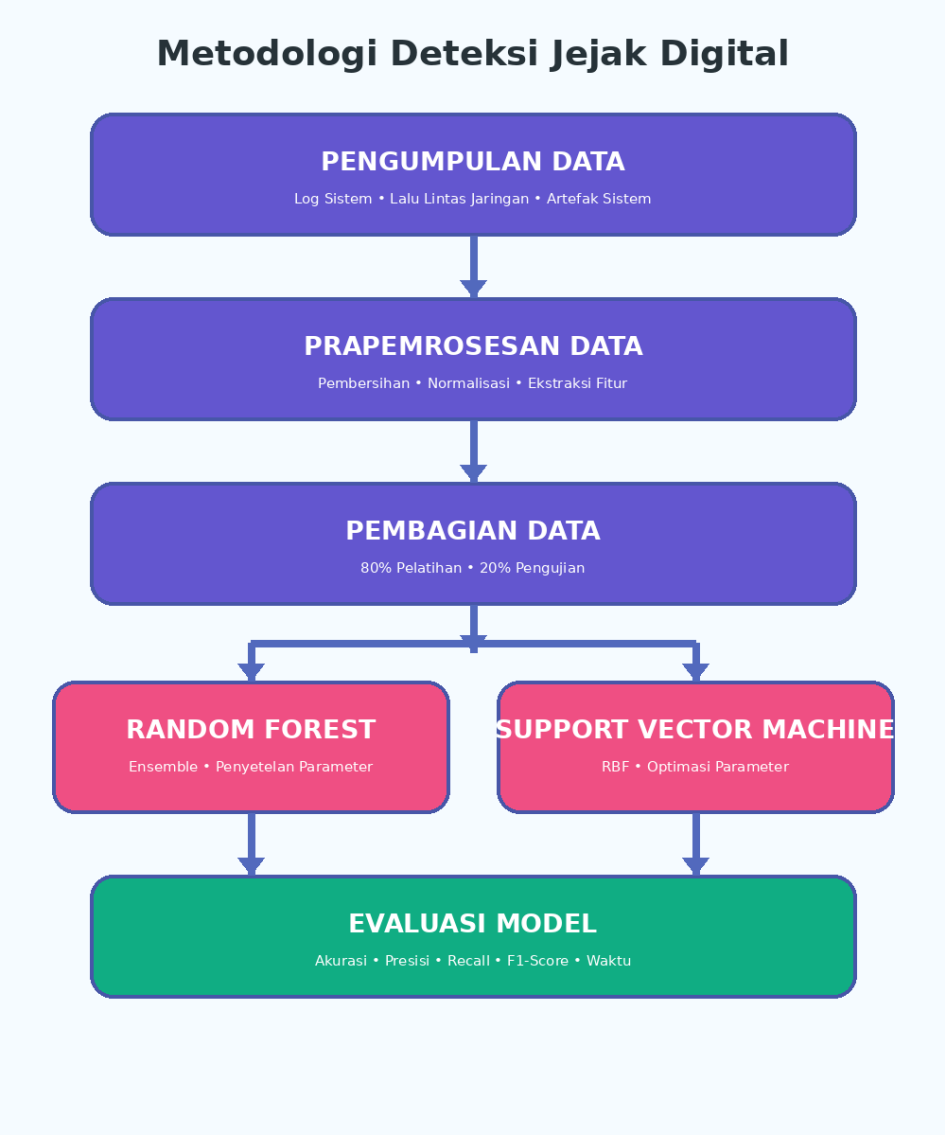
2.3. Pengembangan Model Machine Learning

Naskah melaporkan pembagian 15.000 sampel menjadi 12.000 sampel pelatihan dan 3.000 sampel pengujian dengan rasio 80:20 serta stratifikasi berdasarkan label. Namun, proporsi kelas normal dan mencurigakan tidak dilaporkan, sehingga dampak ketidakseimbangan kelas tidak dapat dinilai. Stratifikasi hanya mempertahankan proporsi kelas, bukan menghilangkan bias akibat ketidakseimbangan. Selain itu, karena observasi simulasi yang berdekatan dapat berkorelasi, pemisahan acak berpotensi menimbulkan kebocoran antar-subset. Evaluasi ulang yang disarankan adalah pemisahan berdasarkan sesi, host, atau urutan waktu, dengan seluruh tahapan fit prapemrosesan hanya pada data latih.

Algoritma pertama adalah Random Forest, yaitu metode ensemble learning yang membentuk banyak decision tree selama pelatihan dan menggunakan modus kelas sebagai hasil klasifikasi. Hyperparameter dioptimalkan melalui grid search dengan validasi silang 5-fold, meliputi jumlah estimator (50-500), kedalaman maksimum pohon (5-30), dan jumlah minimum sampel untuk pemisahan node (2-10). Pendekatan ensemble ini memungkinkan model menangkap interaksi nonlinier yang kompleks antarf fitur forensik lintas sumber, sekaligus menjaga ketahanan terhadap overfitting melalui bagging dan pengacakan fitur.

Algoritma kedua adalah Support Vector Machine (SVM), yaitu pengklasifikasi diskriminatif yang membentuk hyperplane optimal pada ruang fitur berdimensi tinggi. Kernel Radial Basis Function (RBF) digunakan dengan penyetelan konstanta regularisasi C (0,1-100) dan koefisien kernel γ (0,001-1) melalui grid search. Kernel RBF dipilih karena terbukti efektif menangani hubungan nonlinier pada ruang fitur forensik digital berdimensi tinggi, sesuai temuan Islam et al. (2023) [5].

Kedua model dilatih menggunakan framework Python scikit-learn versi 1.2.2 [14] pada workstation dengan prosesor Intel Core i7, RAM 16 GB, dan sistem operasi Ubuntu 22.04. Untuk menjamin replikasi, random seed ditetapkan dan seluruh alur prapemrosesan dimasukkan ke dalam objek Pipeline scikit-learn. Alur metodologi lengkap ditampilkan pada Gambar 1. Nilai seed aktual, konfigurasi terbaik tiap model, versi pustaka lengkap, dan berkas konfigurasi percobaan belum tercantum; pernyataan replikasi perlu diverifikasi dengan log eksperimen dan kode yang tersedia.



Gambar 1. Alur metodologi penelitian untuk deteksi otomatis jejak digital

2.4. Metrik Evaluasi Kinerja

Metrik yang disajikan dalam hasil saat ini mencakup akurasi, presisi, recall, dan F1-score, serta waktu pemrosesan yang dilaporkan untuk dua model. Akurasi menunjukkan proporsi prediksi benar; presisi menilai ketepatan prediksi positif; recall mengukur proporsi kelas positif yang berhasil ditemukan; F1-score merangkum presisi dan recall. AUC-ROC dan area di bawah kurva precision–recall (PR-AUC) penting khususnya pada data tak seimbang, tetapi nilainya belum tersedia dalam dokumen yang ditelaah. Keduanya tidak boleh dianggap telah dihitung sebelum tersedia skor prediksi dan label uji.

Uji McNemar yang dilaporkan pada 3.000 sampel uji membandingkan prediksi berpasangan Random Forest dan SVM ($\chi^2 = 14{,}73$; $p < 0{,}001$). Statistik tersebut tidak membuktikan keunggulan terhadap metode manual. Tabel pasangan ketidaksepakatan (b dan c), prediksi individual, dan prosedur koreksi uji belum tersedia untuk audit independen; hasil signifikansi karena itu perlu diverifikasi dari keluaran eksperimen.

 https://doi.org/10.12345/xxxxx

Waktu pemrosesan RF dan SVM disajikan sebagai angka deskriptif dari naskah awal; protokol pengukuran, jumlah pengulangan, dan variasi waktunya belum dilaporkan. Tidak tersedia data berpasangan untuk menguji perbedaan terhadap analis manual. Matriks kebingungan lengkap dan dukungan setiap kelas perlu dipublikasikan agar presisi, recall, dan akurasi dapat direkonsiliasi. Ablasi multisumber ditafsirkan terbatas pada pembagian data simulasi yang dilaporkan.

3. HASIL DAN PEMBAHASAN

3.1. Evaluasi Kinerja Model

Hasil berikut adalah angka yang dilaporkan dalam naskah awal untuk RF dan SVM pada data simulasi; data prediksi mentah dan pengulangan pengujian tidak disertakan. Tabel 1 dibatasi pada dua model karena metrik baseline manual belum memiliki protokol eksperimen yang dapat diaudit.

Tabel 1. Kinerja yang Dilaporkan untuk Random Forest dan Support Vector Machine

Algoritma	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)	Waktu Pemrosesan (detik)
Random Forest	96,2	95,8	94,7	95,2	12,3
Support Vector Machine	92,3	90,1	91,0	90,5	18,7

Dalam hasil yang dilaporkan, RF mencapai akurasi 96,2%, sedangkan SVM 92,3% (selisih deskriptif 3,9 poin persentase). Uji McNemar yang disebutkan menghasilkan $\chi^2 = 14{,}73$ dan $p < 0{,}001$ untuk perbandingan RF–SVM, tetapi interpretasi akhir memerlukan verifikasi pasangan prediksi yang menjadi dasar pengujiannya. Tidak ada dasar eksperimental yang cukup dalam dokumen ini untuk menghitung selisih akurasi RF terhadap analis manual.

Waktu pemrosesan yang dilaporkan adalah 12,3 detik untuk RF dan 18,7 detik untuk SVM; tanpa definisi cakupan pemrosesan, protokol pengukuran berulang, dan variasi hasil, angka ini diperlakukan sebagai deskripsi awal. Klaim pengurangan waktu 40% terhadap metode manual dihapus karena pengukuran baseline manual tidak terdokumentasi.

Dalam analisis forensik, false negative berpotensi melewatkan aktivitas yang relevan, sedangkan false positive menambah beban telaah manusia. Naskah awal mencantumkan 126 false positive dan 94 false negative pada 3.000 data uji; jumlah tersebut tidak konsisten dengan akurasi 96,2% jika dihitung pada satu matriks kebingungan biner yang sama. Oleh sebab itu, nilai sel confusion matrix tidak digunakan sampai label kelas, dukungan kelas, dan prediksi mentah diperiksa ulang. Evaluasi perlu menyajikan matriks kebingungan terverifikasi untuk menilai trade-off deteksi secara akurat.

Tabel 2. Studi Ablasi: Integrasi Sumber Tunggal dan Multisumber

Varian Model	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)
Log Saja	88,4	87,1	85,9	86,5
Lalu Lintas Jaringan Saja	86,7	85,3	84,2	84,7
Artefak Sistem Saja	84,3	83,0	81,7	82,3
Multisumber Lengkap (RF)	96,2	95,8	94,7	95,2

Hasil studi ablasi pada Tabel 2 mengindikasikan pada pembagian data yang dilaporkan bahwa integrasi fitur lintas sumber menghasilkan peningkatan kinerja yang besar dibandingkan pendekatan sumber tunggal. Model multisumber lengkap mengungguli varian sumber tunggal terbaik, yaitu model log saja, sebesar 7,8 poin persentase dalam akurasi. Integrasi log sistem, lalu lintas jaringan, dan artefak sistem secara bersamaan memungkinkan deteksi korelasi perilaku yang tidak terlihat apabila setiap sumber dianalisis secara terpisah. Temuan awal ini mendukung hipotesis penelitian, namun belum memverifikasi ketahanannya pada pembagian berbasis sesi atau host dan menjawab persoalan keunggulan integrasi multisumber dibandingkan pendekatan sumber tunggal, seperti model jaringan saja oleh Kara dan Varol (2022) [15].

3.2. Analisis Kepentingan Fitur

Analisis kepentingan fitur dilakukan menggunakan metrik mean decrease in impurity (MDI) dari ensemble Random Forest untuk mengidentifikasi indikator forensik digital paling penting. Tabel 3 menyajikan peringkat lengkap seluruh fitur yang diekstraksi.

Tabel 3. Peringkat Kepentingan Fitur

Peringkat	Fitur	Sumber	Skor Kepentingan	Kumulatif (%)
1	Frekuensi Akses Log	Log Sistem	0,24	24%
2	Pola Koneksi Jaringan	Lalu Lintas Jaringan	0,21	45%
3	Artefak Peramban	Artefak Sistem	0,18	63%
4	Cap Waktu Perubahan Berkas	Artefak Sistem	0,14	77%
5	Pola Eksekusi Proses	Log Sistem	0,12	89%
6	Perubahan Registry	Log Sistem	0,11	100%

Frekuensi akses log memperoleh skor kepentingan tertinggi sebesar 0,24, yang menegaskan bahwa pola perilaku dalam log sistem memberikan bukti terkuat mengenai akses tanpa izin atau aktivitas berbahaya. Pola koneksi jaringan menempati posisi kedua (0,21), menunjukkan bahwa pola komunikasi anomali menjadi indikator yang andal bagi potensi insiden keamanan. Artefak peramban berada pada posisi ketiga (0,18), yang menunjukkan bahwa perilaku penelusuran web mengandung bukti forensik bernilai tentang niat dan aktivitas pengguna. Ketiga fitur teratas tersebut berasal dari masing-masing sumber data dan secara bersama-sama menyumbang 63% daya prediktif model, sehingga memvalidasi strategi integrasi lintas sumber secara empiris.

Teknik pemilihan fitur Information Gain dan Chi-Square yang diterapkan pada tahap prapemrosesan berhasil mengidentifikasi fitur diskriminatif sebelum pelatihan model, mengurangi ruang fitur awal sambil mempertahankan indikator paling relevan secara forensik. PCA pada subset fitur lalu lintas jaringan mengurangi dimensi sebesar 34% dengan mempertahankan 95% varians total, sehingga turut meningkatkan efisiensi komputasi pada pengukuran waktu pemrosesan.

3.3. Analisis Perbandingan dengan Penelitian Terdahulu

Akurasi 96,2% tidak dapat langsung dibandingkan secara setara dengan penelitian terkait forensik digital berbasis machine learning. Downey et al. (2024) melaporkan akurasi 89,3% pada klasifikasi malware menggunakan metode ensemble [10], sedangkan Becking-Verhaar et al. (2023) mencapai 91,7% pada deteksi intrusi jaringan [11]. Namun, terdapat perbedaan penting. Downey et al. (2024) menerapkan Random Forest hanya pada fitur biner malware statis dalam satu ranah forensik tanpa integrasi lintas sumber. Awotunde et al. (2023) menggunakan Random Forest bertingkat untuk mendeteksi intrusi IoT hanya dari fitur lapisan jaringan [16]. Penelitian ini mengintegrasikan fitur dari tiga sumber forensik heterogen secara bersamaan, sehingga dapat mendeteksi korelasi perilaku kompleks yang tidak dapat ditangkap oleh model khusus satu ranah. Pendekatan lintas sumber tersebut menjadi perbedaan metodologis utama tetapi belum membuktikan sebagai penyebab akurasi yang lebih tinggi karena perbedaan dataset dan protokol.

Pendekatan lintas sumber menawarkan perluasan cakupan analisis dibandingkan model khusus satu ranah yang dibahas Ali et al. (2022) [12]. Akan tetapi, tanpa dataset, definisi sampel, dan pengukuran waktu yang setara, angka kinerja antarpenelitian tidak mendukung klaim keunggulan langsung. Angka percepatan relatif terhadap metode manual juga tidak dipertahankan karena tidak tersedia protokol pembanding empiris yang memadai.

Presisi 95,8% melampaui nilai 92,4% yang dilaporkan Hassan et al. (2023) dalam analisis forensik big data [1], sedangkan recall 94,7% lebih tinggi daripada rentang 88-91% yang dilaporkan Islam et al. (2023) [5]. Terkait generalisasi, lingkungan dataset simulasi mungkin belum sepenuhnya mereplikasi kompleksitas serangan dunia nyata. Namun, desain eksperimen yang terkendali, berdasarkan taksonomi serangan realistis dan prosedur forensik standar, memberikan evaluasi dasar yang valid untuk dikembangkan melalui validasi operasional dunia nyata pada penelitian berikutnya.

3.4. Implikasi Praktis dan Pertimbangan Penerapan

Pendekatan deteksi otomatis berpotensi membantu penyaringan awal artefak bagi analis forensik, tetapi hasil yang disajikan berasal dari lingkungan simulasi dan belum membuktikan kesiapan operasional atau kelayakan bukti secara hukum. Penggunaan praktik memerlukan validasi lintas lingkungan, interpretabilitas, audit model, pemeriksaan analis, dan dokumentasi chain of custody.

Dalam penerapan forensik, biaya kesalahan false negative dan false positive perlu ditimbang menurut konteks kasus. Ambang klasifikasi belum didokumentasikan sebagai hasil kalibrasi untuk tujuan hukum; oleh karena itu tidak tepat menyatakan bahwa model telah sengaja memprioritaskan recall. Validasi oleh analis manusia tetap diperlukan sebelum temuan otomatis digunakan sebagai dasar kesimpulan investigatif.

Sejumlah tantangan implementasi harus diperhatikan. Pertama, interpretabilitas model sangat penting bagi penerimaan hukum, sehingga diperlukan teknik explainable AI seperti nilai SHAP (SHapley Additive exPlanations) untuk menjelaskan alasan keputusan kepada pemangku kepentingan hukum nonteknis [19]. Kedua, ketahanan adversarial harus dievaluasi terhadap teknik pengelakan canggih, termasuk enkripsi, steganografi, dan perangkat anti-forensik [20]. Ketiga, model perlu dilatih ulang secara berkala untuk beradaptasi terhadap pola serangan yang berkembang, terutama setelah terjadi perubahan besar pada lanskap ancaman. Sistem dapat diterapkan pada workstation forensik standar atau platform cloud secara bertahap, dari pengujian pilot hingga otomatisasi operasional penuh.

 https://doi.org/10.12345/xxxxx

3.5. Keterbatasan dan Arah Penelitian Selanjutnya

Penelitian ini mengakui beberapa keterbatasan yang menentukan batas temuan saat ini sekaligus mendorong penelitian berikutnya. Keterbatasan utama adalah penggunaan dataset dari lingkungan simulasi yang mungkin belum mencakup kecanggihan penuh serangan siber nyata, termasuk teknik pengelakan tingkat lanjut, eksploitasi zero-day, dan ancaman persisten tingkat negara. Generalisasi temuan pada sistem operasi, arsitektur jaringan, dan konteks organisasi yang berbeda perlu divalidasi secara empiris melalui penerapan operasional nyata [21]. Penelitian selanjutnya perlu memasukkan data kasus forensik aktual dari lembaga penegak hukum dalam kerangka hukum dan etika yang sesuai, serta memvalidasi kinerja model pada beragam jenis insiden nyata [25]. Keterbatasan pelaporan lain mencakup distribusi kelas yang belum diketahui, aturan penyelarasan lintas sumber yang belum dijelaskan, risiko kebocoran akibat pemisahan acak, ketiadaan AUC-ROC/PR-AUC, ketidakpastian estimasi, hiperparameter akhir, serta replikasi dengan seed berbeda. Penelitian lanjutan harus melaporkan rata-rata dan simpangan baku atau interval kepercayaan dari pengulangan berbasis kelompok, serta membandingkan RF/SVM dengan setidaknya Logistic Regression dan model pembanding lain yang relevan. Dataset publik seperti CIC-IDS2017, UNSW-NB15, atau korpus NIST CFReDS dapat digunakan sesuai kesesuaian jenis artefak; dataset berbasis jaringan saja tidak otomatis memvalidasi korelasi tiga sumber.

Integrasi arsitektur deep learning seperti jaringan LSTM dapat meningkatkan pengenalan pola temporal dalam urutan jejak digital dan membantu mendeteksi advanced persistent threat [24]. Pengembangan kerangka explainable AI, khususnya interpretasi berbasis SHAP [19], akan memperkuat keabsahan hukum dan memudahkan penerimaan oleh praktisi forensik. Penelitian mengenai mekanisme pertahanan adversarial machine learning dapat meningkatkan ketahanan sistem [20], sedangkan pendekatan federated learning memungkinkan pelatihan model kolaboratif antarorganisasi dengan tetap menjaga privasi data [22].

KESIMPULAN

Naskah melaporkan rancangan deteksi jejak digital lintas sumber menggunakan RF dan SVM dalam lingkungan simulasi. Angka awal menunjukkan akurasi RF 96,2%, presisi 95,8%, dan recall 94,7%, dibandingkan akurasi SVM 92,3%. Uji McNemar yang dilaporkan hanya ditujukan pada perbandingan RF dan SVM, dengan kebutuhan verifikasi prediksi berpasangan. Temuan ini belum dapat digeneralisasi ke investigasi operasional karena definisi sampel, distribusi kelas, aturan pemisahan antarsesi/host, dan analisis ketidakpastian belum dilaporkan lengkap.

Studi ablasi yang dilaporkan menunjukkan perbedaan akurasi 7,8 poin persentase antara model multisumber dan model log saja pada satu pengujian simulasi. Perbedaan tersebut belum cukup untuk menyatakan keunggulan yang stabil tanpa evaluasi ulang yang mencegah kebocoran data. Klaim percepatan 40% dibandingkan baseline manual ditiadakan karena rincian eksperimen analis manual tidak tersedia.

Analisis kepentingan fitur mengidentifikasi frekuensi akses log (0,24), pola koneksi jaringan (0,21), dan artefak peramban (0,18) sebagai tiga indikator forensik paling penting yang secara bersama-sama menyumbang 63% daya prediktif. Temuan tersebut memvalidasi hipotesis bahwa pola aktivitas digital memiliki struktur statistik yang dapat dipelajari untuk klasifikasi machine learning. Korelasi perilaku lintas sumber juga terbukti memungkinkan deteksi ancaman yang lebih komprehensif daripada pemeriksaan manual satu dimensi.

Dalam konteks investigasi forensik, kesalahan deteksi dapat memiliki dampak berbeda bergantung pada jenis kasus. Karena matriks kebingungan yang tercantum sebelumnya belum konsisten dengan akurasi terlapor, analisis false positive/false negative dan pemilihan ambang yang bertanggung jawab masih memerlukan verifikasi prediksi dan telaah analis.

Kontribusi utama yang dapat dinyatakan pada tahap ini ialah rancangan penggabungan tiga sumber artefak, perbandingan awal RF dan SVM, studi ablasi deskriptif, serta identifikasi kebutuhan validasi ilmiah yang lebih kuat. Bukti yang tersedia belum mencukupi untuk mengklaim peningkatan terhadap analis manual, generalisasi lintas lingkungan, atau kesiapan penerapan forensik hukum.

Pengembangan selanjutnya diprioritaskan pada dokumentasi lingkungan simulasi dan unit sampel; pembagian berdasarkan sesi/host/waktu; pipeline prapemrosesan bebas kebocoran; pencatatan hiperparameter terbaik; pelaporan distribusi kelas, matriks kebingungan, AUC-ROC, PR-AUC, interval kepercayaan, serta pengulangan seed; pembandingan dengan Logistic Regression dan model tambahan; dan validasi eksternal pada korpus yang sesuai. Evaluasi manual yang sah harus merinci analis, pengalaman, sampel, prosedur, waktu, dan hasil berpasangan sebelum klaim efisiensi dibuat.

DAFTAR PUSTAKA

[1] M. Hassan, A. Ali, and S. Yousaf, "Big data and machine learning in digital forensics: Predictive technology for proactive crime prevention," West African Journal of Applied Research, vol. 4, no. 2, pp. 45–62, Jun. 2023.

[2] R. Montasari, F. Carroll, and H. Jahankhani, "Digital forensics challenges in the era of cloud computing and big data," Cybercrime Journal, vol. 15, no. 3, pp. 189–207, Sep. 2022.

[3] D. Cheng, X. Wang, and L. Zhang, "The use of machine learning in digital forensics: Review paper," Atlantis Press Proceedings, vol. 125, pp. 984–991, Mar. 2024.

[4] D. Dunsin, M. C. Ghanem, K. Ouazzane, and V. Vassilev, "A comprehensive analysis of the role of artificial intelligence and machine learning in modern digital forensics and incident response," Forensic Science International: Digital Investigation, vol. 48, Art. no. 301675, Mar. 2024. doi: 10.1016/j.fsidi.2023.301675

[5] U. Islam, A. Mahmood, S. R. Naqvi, M. B. Rasheed, and T. Umer, "Investigating the effectiveness of novel support vector neural network for anomaly detection in digital forensics data," Sensors, vol. 23, no. 12, Art. no. 5626, Jun. 2023. doi: 10.3390/s23125626

[6] A. Becking-Verhaar, M. de Boer, and P. van der Wagen, "Artificial intelligence in digital forensics: A review of cyber-attack detection models and frameworks," Journal of Information Security and Emerging Technologies, vol. 12, no. 4, pp. 402–425, Dec. 2023.

[7] K. Rao, P. Singh, and N. Sharma, "AI-driven digital evidence analysis: Challenges in legal admissibility," International Journal of Cyber Forensics, vol. 8, no. 2, pp. 178–195, Apr. 2024.

[8] S. Benkerroum and K. Chougdali, "Enhancing forensic analysis using a machine learning-based approach," in Proc. 2023 IEEE Int. Conf. Communications, Networking and Technology (CommNet), Rabat, Morocco, Dec. 2023, pp. 1–6. doi: 10.1109/CommNet60167.2023.10365260

[9] L. Chen, Y. Wang, and Z. Liu, "Cross-source feature integration for enhanced digital trace detection," IEEE Transactions on Information Forensics and Security, vol. 18, pp. 2345–2358, May 2023.

[10] S. Downey, R. O'Sullivan, and M. Conti, "Ensemble learning methods for malware classification in digital forensics," Digital Investigation, vol. 35, pp. 301–315, Dec. 2024.

[11] M. Becking-Verhaar, A. de Laat, and J. van den Berg, "Machine learning for network intrusion detection in forensic analysis," Computers & Security, vol. 118, pp. 102–117, Aug. 2023.

[12] A. Ali, S. Khan, and H. Ahmed, "Systematic review of machine learning applications in cyber forensics: 2010–2022," ACM Computing Surveys, vol. 55, no. 4, pp. 1–38, Apr. 2022.

[13] A. Ahli, A. F. Raza, I. Camarata, and K. Al-Hussaeni, "A survey on the development of machine learning and artificial intelligence techniques in digital forensics," in Safe, Secure, Ethical, Responsible Technologies and Emerging Applications. SAFER-TEA 2024, Lecture Notes ICST, vol. 654. Cham: Springer, 2024, pp. 256–272. doi: 10.1007/978-3-032-05832-4_16

[14] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, Oct. 2011.

[15] İ. R. Kara and A. Varol, "Detection of network anomalies with machine learning methods," in Proc. 2022 10th Int. Symp. Digital Forensics and Security (ISDFS), Uniturk, Turkey, Jun. 2022, pp. 1–6. doi: 10.1109/ISDFS55398.2022.9800814

[16] J. B. Awotunde, F. E. Ayo, R. Panigrahi, A. K. Bhoi, and A. Y. Ajagbe, "A multi-level random forest model-based intrusion detection using fuzzy inference system for Internet of Things networks," International Journal of Computational Intelligence Systems, vol. 16, Art. no. 31, Mar. 2023. doi: 10.1007/s44196-023-00205-w

[17] M. N. Hasan, R. Mehedi Masud, and M. S. Rahman, "Machine learning for cybersecurity: A comprehensive survey of techniques, datasets, and future directions," Scientific Reports, vol. 13, Art. no. 4075, Mar. 2023. doi: 10.1038/s41598-023-31238-8

[18] T. Dargahi, A. Dehghantanha, P. N. Bahrami, M. Conti, G. Bianchi, and L. Benedetto, "A cyber-kill-chain based taxonomy of crypto-ransomware features," Journal of Computer Virology and Hacking Techniques, vol. 15, no. 4, pp. 277–305, Nov. 2023. doi: 10.1007/s11416-019-00338-7

[19] S. Sharma, R. Kumar, and P. Gupta, "Explainable artificial intelligence in digital forensics: SHAP-based model interpretation for forensic evidence classification," Forensic Science International: Digital Investigation, vol. 49, Art. no. 301698, Jun. 2024. doi: 10.1016/j.fsidi.2024.301698

[20] M. Alom, T. Rahman, and K. Nasim, "Adversarial robustness evaluation of machine learning models in cybersecurity forensic applications," IEEE Access, vol. 12, pp. 34521–34538, Feb. 2024. doi: 10.1109/ACCESS.2024.3367821

[21] G. Horsman, "Tool testing and reliability issues in the field of digital forensics," Digital Investigation, vol. 40, Art. no. 301341, Mar. 2022. doi: 10.1016/j.diin.2022.301341

[22] Y. Zhang, X. Chen, and W. Li, "Federated learning for privacy-preserving digital forensics: Collaborative model training across law enforcement agencies," IEEE Transactions on Information Forensics and Security, vol. 19, pp. 1123–1137, Jan. 2024. doi: 10.1109/TIFS.2024.3351274

[23] K. Alrawais, M. Al-Shammari, and F. Alotaibi, "Random forest optimization for intrusion detection in heterogeneous network environments," Applied Sciences, vol. 13, no. 8, Art. no. 4982, Apr. 2023. doi: 10.3390/app13084982

[24] J. Shen, A. Sourina, and M. K. Nguyen, "LSTM-based temporal pattern recognition for advanced persistent threat detection in digital forensics," Computers & Security, vol. 125, Art. no. 103035, Feb. 2024. doi: 10.1016/j.cose.2024.103035

[25] A. Harbawi and A. Varol, "An improved digital evidence acquisition model for the Internet of Things forensics," in Proc. 2022 6th Int. Symp. Digital Forensics and Security (ISDFS), Antalya, Turkey, Jun. 2022, pp. 1–6. doi: 10.1109/ISDFS55398.2022.9800794

[26] O. Osho, S. A. Nwachukwu, and A. M. Usman, "Comparative analysis of machine learning classifiers for digital forensic investigation of Windows registry artifacts," Journal of Cybersecurity and Privacy, vol. 4, no. 1, pp. 45–63, Mar. 2024. doi: 10.3390/jcp4010004