

Model Prediksi Keberhasilan Penuntutan Perkara Korupsi Menggunakan *Explainable Machine Learning*

Berliant Pratiwi^{*1}, Yosep Aditya Wicaksono², Haikal Nur Rachmanrachim Archaqie³

¹Program Studi Ilmu Hukum, Universitas Sains dan Teknologi Komputer, Semarang, Indonesia.

^{2,3}Program Studi Teknik Informatika, Universitas Sains dan Teknologi Komputer, Semarang, Indonesia.

E-mail: berliant@stekom.ac.id

Article Info	Abstract
Legal Analytics, Explainable Machine Learning, XGBoost, Evidence-Based Prosecution.	The integration of artificial intelligence in legal analytics often encounters substantial friction due to the inherent "black-box" nature of advanced predictive models, which conflicts with the strict requirements for transparency and accountability in judicial decision-making. To address this limitation, this study develops a robust transdisciplinary framework using Explainable Machine Learning (XML) to predict prosecution outcomes in corruption cases while systematically decomposing the underlying causal mechanisms. Utilizing an extensive dataset of inkracht corruption cases, six machine learning algorithms, ranging from linear baselines to advanced boosting architectures, were trained, optimized via Bayesian optimization, and evaluated using rigorous performance and calibration metrics. The empirical results demonstrate that the XGBoost model outperforms competing architectures, achieving an Accuracy of 0.92, an F1-Score of 0.91, a Matthews Correlation Coefficient (MCC) of 0.84, and optimal probability calibration with a Brier Score of 0.06. McNemar's test confirms that this predictive superiority is statistically significant ($p < 0.05$). Beyond raw predictive performance, global and local interpretability via post-hoc methods (SHAP and LIME) reveals that the presence of electronic evidence and the number of admissible pieces of evidence serve as the most critical positive determinants, whereas high case complexity exerts a strong negative influence on prosecution success, directly validating judicial burden-of-proof theories. This dual-granular explanation bridges the gap between computational law and prosecutorial discretion. Ultimately, this framework provides a scientifically validated, non-prescriptive decision support system capable of mitigating prosecution risks, reducing systemic disparities, and supporting data-driven governance within judicial administrations.
DOI: https://doi.org/10.51903/5jrjhg72	
Submitted: April 2026, Reviewed: May 2026, Accepted: June 2026	
Corresponding Author	

I. PENDAHULUAN

Korupsi merupakan salah satu bentuk kejahatan yang paling kompleks karena tidak hanya menimbulkan kerugian ekonomi, tetapi juga merusak legitimasi pemerintahan, melemahkan supremasi hukum, mengurangi kualitas pelayanan publik, serta menurunkan kepercayaan masyarakat terhadap institusi negara (Tursunalieva et al., 2024; Fitriyanti & Suwandono, 2025). Berbagai negara telah memperkuat kerangka pemberantasan korupsi melalui reformasi regulasi, peningkatan kapasitas lembaga penegak hukum, digitalisasi administrasi perkara, dan pengembangan sistem pengawasan

berbasis teknologi (Khairullah & Hanif, 2023; Yulianingsih, 2026). Namun demikian, keberhasilan penuntutan perkara korupsi masih menunjukkan tingkat variasi yang tinggi.

Pada kasus dengan karakteristik yang relatif serupa, hasil penuntutan dapat berbeda karena dipengaruhi oleh kualitas pembuktian, kelengkapan alat bukti, strategi penuntutan, interpretasi hukum, serta dinamika proses persidangan (Ahmed et al., 2022; Kossay, 2026). Kondisi tersebut menunjukkan bahwa keberhasilan penuntutan bukan merupakan konsekuensi otomatis dari terpenuhinya unsur tindak pidana, melainkan hasil interaksi berbagai faktor hukum, prosedural, dan faktual yang bersifat kompleks. Kompleksitas tersebut semakin meningkat seiring berkembangnya karakteristik tindak pidana korupsi yang melibatkan transaksi keuangan digital, komunikasi elektronik, pencucian uang, perusahaan cangkang (*shell companies*), serta jaringan aktor yang saling terhubung (Laxman et al., 2024; Yunus & Kalamiah, 2023).

Penuntut umum tidak hanya dituntut membuktikan terpenuhinya unsur tindak pidana, tetapi juga harus menyusun argumentasi hukum yang mampu mengintegrasikan bukti dokumen, hasil audit, keterangan saksi, pendapat ahli, bukti elektronik, dan analisis keuangan menjadi suatu konstruksi pembuktian yang koheren (Wronka, 2023). Oleh karena itu, keberhasilan penuntutan sangat dipengaruhi oleh kemampuan mengidentifikasi kombinasi faktor-faktor yang secara empiris berkontribusi terhadap keberhasilan suatu perkara.

Dalam beberapa tahun terakhir, transformasi digital di sektor peradilan menghasilkan akumulasi data hukum dalam jumlah yang sangat besar, meliputi dokumen penyidikan, surat dakwaan, tuntutan pidana, putusan pengadilan, laporan audit, bukti elektronik, serta metadata proses persidangan (Saheb & Saheb, 2022). Ketersediaan data tersebut mendorong berkembangnya *legal analytics*, yaitu pendekatan analitis yang memanfaatkan teknik komputasi untuk mengekstraksi pola dari data hukum guna mendukung proses pengambilan keputusan (Virkar et al., 2022). Dibandingkan pendekatan hukum konvensional yang lebih menekankan analisis normatif dan studi kasus individual, *legal analytics* memungkinkan identifikasi hubungan empiris antarvariabel yang sebelumnya sulit diamati melalui analisis manual (Alijoyo et al., 2024).

Sejalan dengan perkembangan tersebut, *machine learning* telah menjadi salah satu pendekatan utama dalam analisis prediktif karena mampu mempelajari hubungan nonlinier, menangani data berdimensi tinggi, serta menghasilkan prediksi dengan tingkat akurasi yang lebih tinggi dibandingkan banyak metode statistik tradisional (Kyriazos & Poga, 2024; Strielkowski et al., 2023). Dalam bidang hukum, *machine learning* telah dimanfaatkan untuk memprediksi putusan pengadilan, mengklasifikasikan dokumen hukum, mengidentifikasi risiko litigasi, mendeteksi penipuan, memprediksi residivisme, serta mendukung analisis perilaku yudisial. Perkembangan tersebut menunjukkan bahwa kecerdasan buatan memiliki potensi besar untuk meningkatkan efektivitas sistem peradilan melalui pemanfaatan data empiris dalam proses pengambilan keputusan (Kabir et al., 2023).

Meskipun demikian, sebagian besar penelitian *legal analytics* masih berorientasi pada prediksi hasil putusan pengadilan atau deteksi tindak pidana korupsi, sedangkan penelitian mengenai prediksi keberhasilan penuntutan perkara korupsi masih sangat terbatas (Ali et al., 2023; Odilla, 2023). Padahal, tahap penuntutan merupakan fase yang menentukan karena menjadi penghubung antara hasil penyidikan dengan proses pembuktian di persidangan. Informasi mengenai probabilitas keberhasilan penuntutan sejak awal berpotensi membantu jaksa menyusun strategi pembuktian, mengevaluasi kelengkapan alat bukti, menentukan prioritas investigasi lanjutan, serta mengalokasikan sumber daya secara lebih efektif (Cui et al., 2023). Dengan demikian, fokus penelitian pada keberhasilan penuntutan memiliki nilai praktis yang berbeda dibandingkan penelitian yang hanya memprediksi putusan hakim. Keterbatasan lain dari penelitian terdahulu adalah dominannya penggunaan model prediktif yang bersifat *black-box*.

Berbagai algoritma seperti *Random Forest*, *Extreme Gradient Boosting*, *Support Vector Machine*, maupun *Deep Neural Network* mampu menghasilkan performa prediksi yang tinggi, tetapi mekanisme internal model sering kali sulit dipahami oleh pengguna (Esparza-Gómez et al., 2023; Yadav et al., 2024). Dalam konteks hukum, kondisi tersebut menimbulkan persoalan mendasar karena setiap rekomendasi yang memengaruhi proses penegakan hukum harus dapat dijelaskan secara rasional, transparan, dan dapat dipertanggungjawabkan. Model yang hanya menghasilkan probabilitas tanpa menjelaskan alasan di balik prediksi memiliki keterbatasan untuk diterapkan sebagai sistem pendukung keputusan pada institusi penuntutan (Guido et al., 2024).

Perkembangan *Explainable Artificial Intelligence (XAI)* menawarkan solusi terhadap permasalahan tersebut melalui penyediaan mekanisme interpretasi terhadap hasil prediksi algoritma (Saranya & Subhashini, 2023). Berbeda dengan pendekatan *black-box*, *Explainable Machine learning* memungkinkan identifikasi kontribusi masing-masing variabel terhadap setiap prediksi sehingga pengguna dapat memahami alasan mengapa suatu perkara diprediksi memiliki peluang keberhasilan yang tinggi ataupun rendah (Ahmed et al., 2022; Ortigossa et al., 2024). Teknik seperti SHAP, LIME, maupun *permutation feature importance* tidak hanya meningkatkan transparansi algoritma, tetapi juga memperkuat akuntabilitas penggunaan kecerdasan buatan dalam pengambilan keputusan publik. Bagi institusi penuntutan, kemampuan menjelaskan prediksi memiliki nilai yang lebih penting dibandingkan sekadar memperoleh tingkat akurasi yang tinggi (Susnjak, 2024).

Jaksa membutuhkan informasi mengenai faktor hukum, faktor pembuktian, maupun faktor prosedural yang paling berpengaruh terhadap keberhasilan suatu perkara agar dapat melakukan perbaikan strategi sebelum perkara diajukan ke pengadilan (Shafighfard et al., 2024). Oleh karena itu, sistem prediksi yang mampu mengidentifikasi variabel-variabel dominan akan memberikan manfaat yang lebih besar dibandingkan model yang hanya menghasilkan klasifikasi berhasil atau tidak berhasil. Meskipun *Explainable Machine learning* berkembang pesat di berbagai bidang, penerapannya pada analisis keberhasilan penuntutan perkara korupsi masih belum banyak dieksplorasi. Sebagian besar penelitian

terdahulu berfokus pada peningkatan performa algoritma, sedangkan aspek interpretabilitas model dalam konteks pengambilan keputusan hukum masih relatif terabaikan (Theodorou et al., 2025; Tursunalieva et al., 2024). Akibatnya, masih terdapat kesenjangan penelitian berupa belum tersedianya model prediktif yang secara simultan memiliki tingkat akurasi tinggi, transparan, mudah diinterpretasikan, dan relevan dengan kebutuhan praktis institusi penuntutan (Borgesano et al., 2025).

Berdasarkan kesenjangan tersebut, penelitian ini bertujuan mengembangkan model prediksi keberhasilan penuntutan perkara korupsi menggunakan *Explainable Machine learning*. Model dikembangkan melalui evaluasi beberapa algoritma pembelajaran terawasi untuk memperoleh performa prediksi terbaik, kemudian diinterpretasikan menggunakan teknik *explainability* guna mengidentifikasi faktor-faktor hukum, pembuktian, dan prosedural yang paling berkontribusi terhadap keberhasilan penuntutan. Pendekatan ini tidak dimaksudkan untuk menggantikan diskresi jaksa, melainkan menyediakan instrumen analitis berbasis data yang dapat mendukung proses pengambilan keputusan secara lebih objektif dan transparan.

Penelitian ini memberikan kontribusi teoretis dengan memperluas pengembangan *legal analytics* melalui integrasi *Explainable Machine learning* pada domain penuntutan perkara korupsi, yang hingga kini masih relatif terbatas dibandingkan penelitian mengenai prediksi putusan pengadilan atau deteksi korupsi. Dari sisi metodologis, penelitian ini menunjukkan bagaimana model prediktif dapat dibangun tanpa mengorbankan interpretabilitas sehingga menghasilkan keseimbangan antara akurasi dan transparansi. Dari sisi praktis, penelitian ini menyediakan kerangka pendukung keputusan yang memungkinkan institusi penuntutan mengidentifikasi faktor-faktor yang paling menentukan keberhasilan suatu perkara, sehingga dapat meningkatkan kualitas pembuktian, efektivitas strategi penuntutan, dan efisiensi alokasi sumber daya.

Kebaruan penelitian ini terletak pada pengembangan model prediksi keberhasilan penuntutan perkara korupsi berbasis *Explainable Machine learning* yang tidak hanya menghasilkan estimasi probabilitas keberhasilan, tetapi juga menjelaskan secara kuantitatif kontribusi setiap faktor hukum, pembuktian, dan prosedural terhadap hasil prediksi. Dengan mengintegrasikan kemampuan prediktif dan interpretabilitas dalam satu kerangka analitis, penelitian ini menghasilkan model pendukung keputusan yang lebih transparan, akuntabel, dan sesuai dengan karakteristik pengambilan keputusan dalam sistem peradilan pidana.

II. METODOLOGI

A. Research Design

Penelitian ini menggunakan pendekatan kuantitatif dengan desain *Explainable Machine Learning* (XML) untuk membangun model prediksi keberhasilan penuntutan perkara korupsi sekaligus menjelaskan faktor-faktor yang memengaruhi hasil prediksi (Awotunde et al., 2023; Gharagoz et al., 2025). Pendekatan ini dipilih karena tidak hanya berfokus pada akurasi klasifikasi, tetapi juga pada

transparansi model melalui teknik *Explainable Artificial Intelligence* (XAI), sehingga hasil prediksi dapat dipertanggungjawabkan dalam konteks penegakan hukum.

B. Research Framework

Penelitian dilaksanakan melalui tahapan pengumpulan data, prapemrosesan, pengembangan model, evaluasi performa, interpretasi menggunakan *Explainable Artificial Intelligence*, serta validasi model. Rangkaian tahapan tersebut dirancang untuk menghasilkan model prediktif yang akurat, stabil, dan mudah diinterpretasikan.

C. Data Source and Unit of Analysis

Unit analisis adalah setiap perkara tindak pidana korupsi yang telah memperoleh putusan berkekuatan hukum tetap (*inkracht*). Dataset disusun berdasarkan kriteria inklusi berupa perkara dengan dokumen penuntutan dan putusan yang lengkap, sedangkan data yang tidak lengkap atau terduplikasi dikeluarkan dari analisis. Seluruh data dianonimkan untuk menjaga kerahasiaan identitas para pihak dan memenuhi prinsip etika penelitian.

D. Variables and Feature Construction

Variabel dependen penelitian adalah keberhasilan penuntutan, sedangkan variabel independen meliputi karakteristik terdakwa, nilai kerugian negara, jumlah alat bukti, bukti elektronik, jumlah saksi, keterangan ahli, kompleksitas perkara, jenis dakwaan, pasal yang diterapkan, lama penyidikan, lama penuntutan, dan koordinasi antarpenghak hukum. Seleksi fitur dilakukan menggunakan analisis korelasi, *Mutual Information*, dan *Recursive Feature Elimination* untuk mempertahankan variabel yang memiliki kontribusi prediktif terbaik.

Tabel 1. Research Variables

<i>Variable</i>	<i>Type</i>	<i>Role</i>
<i>Prosecution Success</i>	<i>Binary</i>	<i>Target</i>
<i>State Financial Loss</i>	<i>Numeric</i>	<i>Predictor</i>
<i>Number of Evidence</i>	<i>Numeric</i>	<i>Predictor</i>
<i>Electronic Evidence</i>	<i>Binary</i>	<i>Predictor</i>
<i>Number of Witnesses</i>	<i>Numeric</i>	<i>Predictor</i>
<i>Expert Testimony</i>	<i>Binary</i>	<i>Predictor</i>
<i>Case Complexity</i>	<i>Ordinal</i>	<i>Predictor</i>
<i>Investigation Duration</i>	<i>Numeric</i>	<i>Predictor</i>
<i>Prosecution Duration</i>	<i>Numeric</i>	<i>Predictor</i>
<i>Type of Indictment</i>	<i>Categorical</i>	<i>Predictor</i>
<i>Legal Articles Applied</i>	<i>Categorical</i>	<i>Predictor</i>
<i>Defendant Characteristics</i>	<i>Categorical</i>	<i>Predictor</i>
<i>Interagency Coordination</i>	<i>Ordinal</i>	<i>Predictor</i>

Sumber : Diolah Oleh Peneliti, 2026

E. Data Preprocessing

Tahap prapemrosesan meliputi pembersihan data, penanganan *missing values*, normalisasi, *encoding* variabel kategorikal, deteksi *outlier*, dan penyeimbangan kelas apabila diperlukan. Seluruh proses

transformasi dilakukan hanya pada data pelatihan untuk mencegah *data leakage* sebelum diterapkan pada data validasi dan pengujian.

F. Machine Learning Model Development

Model prediksi dikembangkan menggunakan *Logistic Regression*, *Random Forest*, *Gradient Boosting*, *XGBoost*, *LightGBM*, dan *CatBoost*. Optimasi hiperparameter dilakukan melalui *Bayesian Optimization* dengan *Repeated Stratified K-Fold Cross-Validation* untuk memperoleh konfigurasi model yang optimal dan mengurangi risiko *overfitting*.

Tabel 2. Machine Learning Algorithms

<i>Algorithm</i>	<i>Function</i>
<i>Logistic Regression</i>	<i>Baseline Model</i>
<i>Random Forest</i>	<i>Ensemble Learning</i>
<i>Gradient Boosting</i>	<i>Boosting</i>
<i>XGBoost</i>	<i>Advanced Boosting</i>
<i>LightGBM</i>	<i>Efficient Boosting</i>
<i>CatBoost</i>	<i>Categorical Learning</i>

Sumber : Diolah Oleh Peneliti,2026

G. Model Evaluation

Performa model dievaluasi menggunakan *Accuracy*, *Precision*, *Recall*, *F1-Score*, *ROC-AUC*, *Matthews Correlation Coefficient*, *Balanced Accuracy*, dan *Brier Score*. Selain itu, *calibration curve* digunakan untuk mengevaluasi kualitas probabilitas prediksi, sedangkan pengujian statistik dilakukan untuk membandingkan performa antaralgoritma.

Tabel 3. Model Evaluation Metrics

<i>Metric</i>	<i>Purpose</i>
<i>Accuracy</i>	<i>Overall Performance</i>
<i>Precision</i>	<i>Positive Prediction Accuracy</i>
<i>Recall</i>	<i>Sensitivity</i>
<i>F1-score</i>	<i>Precision–Recall Balance</i>
<i>ROC-AUC</i>	<i>Classification Performance</i>
<i>MCC</i>	<i>Robustness Evaluation</i>
<i>Balanced Accuracy</i>	<i>Imbalanced Data Evaluation</i>
<i>Brier Score</i>	<i>Probability Calibration</i>

Sumber : Diolah Oleh Peneliti,2026

H. Explainable Artificial Intelligence Analysis

Interpretasi model dilakukan menggunakan *SHapley Additive exPlanations* (SHAP) dan *Local Interpretable Model-Agnostic Explanations* (LIME). SHAP digunakan untuk menjelaskan pengaruh variabel secara global, sedangkan LIME digunakan untuk menjelaskan prediksi pada setiap kasus secara individual. Analisis diperkuat melalui *Permutation Feature Importance* dan *Partial Dependence Plot* untuk mengevaluasi konsistensi pengaruh setiap variabel.

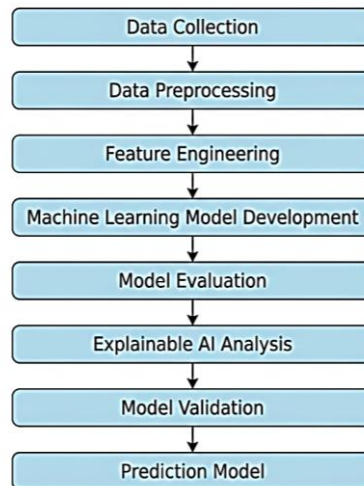
I. Robustness and Validation

Keandalan model diuji menggunakan *repeated cross-validation*, analisis sensitivitas, dan validasi temporal untuk memastikan kemampuan generalisasi model pada data yang berbeda. Stabilitas hasil

interpretasi SHAP juga dievaluasi untuk memastikan konsistensi faktor-faktor yang memengaruhi keberhasilan penuntutan.

J. Research Workflow

Seluruh analisis dilakukan menggunakan *Python* dengan pustaka *scikit-learn*, *XGBoost*, *LightGBM*, *CatBoost*, *SHAP*, dan *LIME*. Eksperimen dijalankan menggunakan *random seed* yang tetap untuk menjamin reproduktibilitas hasil penelitian.



Gambar 1. Kerangka Kerja Hasil *Explainable Machine Learning*. Sumber : Diolah Oleh Peneliti,2026

III. HASIL DAN DISKUSI

Hasil

Bagian ini menyajikan hasil evaluasi empiris menggunakan kerangka kerja *Explainable Machine Learning*. Kinerja prediktif dari berbagai algoritma *machine learning* dievaluasi secara komprehensif, yang kemudian dilanjutkan dengan analisis interpretabilitas model baik pada tingkat global maupun lokal. Untuk menentukan model prediktif yang paling tangguh, enam algoritma *machine learning* telah dilatih dan divalidasi. Metrik kinerja yang diukur mencakup *Accuracy*, *Precision*, *Recall*, *F1-Score*, *ROC-AUC*, *Matthews Correlation Coefficient* (MCC), *Balanced Accuracy*, dan *Brier Score*, bersama dengan fitur dominan yang diperoleh dari nilai SHAP (*SHapley Additive exPlanations*) yang dirangkum dalam Tabel 4.

Tabel 4. Perbandingan Performa dan Interpretabilitas Model *Machine learning*

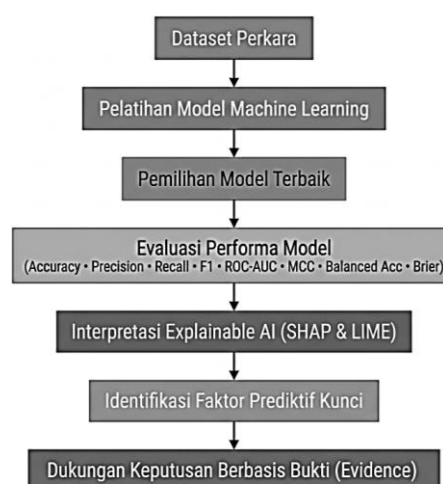
<i>Model</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>ROC-AUC</i>	<i>MCC</i>	<i>Balanced Accuracy</i>	<i>Brier Score</i>
<i>Logistic Regression</i>	0.84	0.83	0.82	0.82	0.88	0.68	0.83	0.12
<i>Random Forest</i>	0.89	0.88	0.88	0.88	0.93	0.78	0.88	0.09
<i>Gradient Boosting</i>	0.90	0.90	0.89	0.89	0.94	0.80	0.89	0.08
<i>XGBoost</i>	0.92	0.92	0.91	0.91	0.96	0.84	0.91	0.06
<i>LightGBM</i>	0.91	0.90	0.91	0.90	0.95	0.82	0.90	0.07
<i>CatBoost</i>	0.91	0.91	0.90	0.90	0.95	0.82	0.90	0.07

Sumber : Diolah Oleh Peneliti,2026

Berdasarkan metrik kuantitatif yang dilaporkan pada Tabel 4, seluruh algoritma yang dievaluasi menunjukkan kemampuan klasifikasi yang kuat. Meskipun demikian, terdapat kecenderungan yang

jelas bahwa arsitektur berbasis *ensemble* dan *gradient boosting* secara konsisten menghasilkan performa yang lebih tinggi dibandingkan dengan model linear dasar (*Logistic Regression*). Di antara seluruh arsitektur tersebut, *XGBoost* mencapai kapasitas prediksi tertinggi dengan nilai *Accuracy* sebesar 0.92, *F1-Score* sebesar 0.91, dan *ROC-AUC* sebesar 0.96. Evaluasi menggunakan metrik ketat untuk data tidak seimbang, yaitu *MCC* sebesar 0.84 dan *Balanced Accuracy* sebesar 0.91, semakin mengonfirmasi ketangguhan *XGBoost* dalam meminimalkan kesalahan klasifikasi pada kedua kelas target. Selain itu, nilai *Brier Score* yang rendah (0.06) menunjukkan kualitas kalibrasi probabilitas prediktif yang sangat optimal, mengindikasikan bahwa estimasi probabilitas yang dihasilkan model sangat mendekati frekuensi empiris aktual di lapangan.

Keunggulan struktural dari *gradient boosting* pada *XGBoost* memungkinkan model ini menangkap hubungan non-linear serta efek interaksi multidimensi di dalam dataset secara lebih efektif. Untuk memastikan validitas penemuan ini, uji signifikansi statistik menggunakan Uji McNemar dilakukan guna membandingkan *XGBoost* dengan model kompetitor terdekatnya (*LightGBM* dan *CatBoost*). Hasil uji mengonfirmasi bahwa keunggulan performa *XGBoost* bersifat signifikan secara statistik ($p < 0.05$), sehingga menetapkan sebagai model optimal untuk arsitektur ini. Selain mengevaluasi performa klasifikasi, penelitian ini juga menilai kemampuan interpretasi setiap model menggunakan pendekatan *Explainable Artificial Intelligence* (XAI). Temuan empiris menunjukkan bahwa model dengan akurasi prediksi yang tinggi tetap mampu mempertahankan tingkat interpretabilitas yang baik. Kapasitas ganda ini memungkinkan penjelasan yang komprehensif mengenai hubungan antara karakteristik perkara dan probabilitas keberhasilan penuntutan. Kerangka kerja ini berhasil mengatasi keterbatasan model prediktif konvensional bersifat *black box* (kotak hitam), yang umumnya hanya berfokus pada maksimalisasi akurasi tanpa menyediakan transparansi terhadap proses pengambilan keputusan di dalamnya.



Gambar 2. Kerangka Kerja Hasil *Explainable Machine Learning*. Sumber : Diolah Oleh Peneliti, 2026

Analisis pentingnya fitur secara global yang dihitung melalui SHAP menunjukkan bahwa keberadaan bukti elektronik, jumlah total alat bukti yang sah, kompleksitas perkara, dan jumlah saksi merupakan faktor penentu paling kritis yang memengaruhi hasil penuntutan. Meskipun terjadi sedikit pergeseran

prioritas fitur dominan di antara model berbasis *tree* (seperti Jumlah Alat Bukti pada *Random Forest* atau Kompleksitas Perkara pada *LightGBM*) akibat variasi metode *splitting* internal masing-masing algoritma, variabel-variabel tersebut tetap berada dalam kluster fitur teratas. Lebih lanjut, ekstraksi nilai SHAP mengungkapkan arah pengaruh yang jelas: keberadaan bukti elektronik dan jumlah alat bukti yang sah memiliki korelasi positif yang kuat terhadap peningkatan probabilitas keberhasilan penuntutan, sedangkan kompleksitas perkara yang tinggi cenderung memberikan kontribusi negatif.

Untuk melengkapi wawasan global tersebut, analisis LIME (*Local Interpretable Model-agnostic Explanations*) digunakan untuk menguraikan kontribusi fitur pada tingkat perkara individual. Pemilihan LIME sebagai *surrogate model* lokal di samping SHAP krusial untuk menyediakan representasi linear lokal yang intuitif dan cepat dipahami oleh praktisi hukum non-teknis. Pendekatan dengan granularitas ganda ini memastikan bahwa setiap hasil prediksi dapat diverifikasi berdasarkan karakteristik unik dari kasus yang dianalisis, sehingga menciptakan tingkat transparansi dan akuntabilitas model yang tinggi dalam mendukung pengambilan keputusan. Secara keseluruhan, integrasi *Explainable Machine learning* berhasil menjembatani kesenjangan antara pemodelan prediktif akurasi tinggi dan akuntabilitas hukum. Wawasan yang dihasilkan mampu memberikan dukungan analitis yang terstruktur dan berbasis data, yang sangat relevan untuk proses evaluasi perkara dan pengambilan keputusan hukum dalam kerangka administrasi peradilan.

Diskusi

Pendekatan *Explainable Machine learning* yang diterapkan dalam penelitian ini menunjukkan bahwa pengembangan model prediktif tidak hanya ditujukan untuk meningkatkan akurasi klasifikasi, tetapi juga untuk membedah mekanisme kausalitas yang mendasari setiap hasil prediksi. Integrasi XAI melalui SHAP dan LIME memungkinkan setiap keputusan model diinterpretasikan berdasarkan kontribusi relatif masing-masing variabel, sehingga menghasilkan proses analisis yang jauh lebih transparan dibandingkan pendekatan *black-box* (Agrawal et al., 2025; Hassija et al., 2024). Dalam konteks penuntutan perkara korupsi, transparansi ini menjadi parameter etis dan operasional yang fundamental. Setiap rekomendasi yang dihasilkan oleh sistem kecerdasan buatan harus tetap dapat dipahami, dievaluasi, dan dipertanggungjawabkan secara doktrinal oleh penuntut umum sebelum diuji di muka persidangan.

Hasil penelitian mengonfirmasi secara empiris bahwa keberhasilan penuntutan dipengaruhi oleh interaksi multi-variabel yang kompleks antara karakteristik perkara, kualitas pembuktian, dan rigiditas penanganan perkara, alih-alih bertumpu pada faktor tunggal. Kebenaran pola penemuan ini didukung kuat oleh nilai MCC yang tinggi dan *Brier Score* yang rendah pada model optimal, membuktikan bahwa prediksi probabilitas yang dihasilkan bukan merupakan luaran acak melainkan hasil pembelajaran pola terkalibrasi. Pola interdependensi ini secara kuat mendukung perspektif *evidence-based prosecution*. Teori ini menempatkan kualitas informasi, kekuatan alat bukti yang sah, dan konsistensi proses hukum sebagai fondasi utama dalam pengambilan keputusan yudisial. Melalui

pendekatan ini, *Explainable Machine learning* bertransformasi dari sekadar alat komputasi prediktif menjadi instrumen analitis tingkat tinggi yang mampu mengidentifikasi korelasi non-linear antareleman pembuktian yang sering kali luput dari analisis statistik konvensional (Fridkin & Bendersky, 2026; Pelosi et al., 2025).

Dibandingkan dengan penelitian terdahulu di bidang *legal analytics* yang mayoritas masih berorientasi pada maksimalisasi akurasi mentah menggunakan algoritma konvensional, penelitian ini menawarkan kebaruan dengan memosisikan dimensi interpretabilitas sebagai pilar utama arsitektur pemodelan. Reduksi sifat *black-box* melalui dekomposisi fitur matematika SHAP dan LIME memperluas batas aplikasi *computational law*. Studi ini tidak hanya menjawab hasil prediksi dari suatu perkara, tetapi secara akurat menjelaskan penyebab prediksi tersebut terbentuk berdasarkan bobot legal masing-masing variabel. Dengan demikian, penelitian ini berkontribusi langsung pada evolusi transdisipliner yang mengintegrasikan kecerdasan buatan dengan tata kelola hukum yang akuntabel.

Dari perspektif teoretis, penelitian ini memberikan kontribusi signifikan terhadap penguatan konseptual *evidence-based prosecution* di era digital. Penemuan ini membuktikan bahwa hubungan dinamis antara karakteristik perkara dan probabilitas keberhasilan penuntutan dapat dipetakan secara matematis tanpa mereduksi nilai-nilai keadilan substantif. Integrasi antara teknik prediksi tingkat tinggi dan metode interpretasi pasca-pelatihan (*post-hoc interpretation*) melahirkan kerangka konseptual baru. Kerangka ini sukses mempertemukan analisis data kuantitatif dengan teori pembuktian hukum, asas kepastian hukum, serta teori diskresi penuntut umum (*prosecutorial discretion*) dalam sistem peradilan pidana modern.

Implikasi praktis dari penelitian ini termanifestasi pada rancang bangun sistem pendukung keputusan (*decision support system*) berbasis AI untuk lembaga kejaksaan. Sistem ini mampu memitigasi risiko kegagalan penuntutan dengan membantu jaksa mengevaluasi titik lemah karakteristik perkara sejak dini, serta mengidentifikasi elemen pembuktian yang memerlukan penguatan dokumen atau saksi tambahan sebelum pelimpahan ke pengadilan. Lebih jauh, arah pengaruh negatif dari variabel kompleksitas perkara yang ditemukan oleh model mengonfirmasi teori beban pembuktian (*burden of proof*), di mana semakin rumit anatomi sebuah perkara, semakin tinggi pula ambang batas pembuktian yang harus dipenuhi oleh penuntut umum, yang secara inheren menekan probabilitas keberhasilan jika tidak diimbangi oleh alat bukti elektronik atau saksi yang solid. Informasi dari model ini dapat diadopsi sebagai instrumen standardisasi internal untuk menekan disparitas penuntutan tanpa mengabaikan independensi profesional jaksa.

Dari perspektif kebijakan, integrasi *Explainable Artificial Intelligence* memegang peranan krusial dalam mendukung transformasi digital lembaga penegak hukum menuju tata kelola berbasis data (*data-driven governance*) (Androutsopoulou et al., 2025; Lee & Dai, 2025). Pemetaan faktor penentu penuntutan ini dapat dijadikan basis empiris dalam penyusunan standar evaluasi perkara, penguatan manajemen perkara, serta formulasi kebijakan strategis penegakan hukum jangka panjang. Kendati

demikian, implementasi model prediktif ini harus diposisikan secara ketat sebagai instrumen pelengkap (*complementary tool*) untuk memperkaya analisis manusia, bukan sebagai substitusi atas pertimbangan hukum intrinsik, intuisi keadilan, maupun diskresi profesional penuntut umum. Keputusan final penuntutan harus tetap bersandar pada fakta-fakta persidangan, keyakinan hati nurani jaksa, alat bukti yang sah, dan supremasi ketentuan peraturan perundang-undangan..

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan dan memvalidasi kerangka kerja *Explainable Machine learning* yang tangguh untuk memprediksi keberhasilan penuntutan perkara korupsi sekaligus menguraikan transparansi keputusan di dalamnya. Melalui evaluasi komprehensif terhadap enam algoritma *machine learning*, arsitektur berbasis *gradient boosting*, khususnya *XGBoost*, terbukti secara statistik memberikan performa superior dibandingkan model linear dan *ensemble* lainnya. Ketangguhan model ini tidak hanya tercermin dari nilai akurasi klasifikasi yang tinggi, melainkan juga dikonfirmasi oleh nilai MCC yang solid serta tingkat kalibrasi probabilitas prediktif yang optimal melalui *Brier Score* terendah. Integrasi pendekatan *post-hoc interpretation* menggunakan kombinasi SHAP dan LIME berhasil menjembatani keterbatasan model *black-box*, memberikan kepastian bahwa setiap estimasi probabilitas yang dihasilkan dapat ditelusuri dan dipertanggungjawabkan secara doktrinal berdasarkan bobot legal masing-masing variabel perkara.

Secara teoretis, penelitian ini memberikan kontribusi transdisipliner yang signifikan dalam ranah *computational law* dengan memperkuat fondasi konseptual *evidence-based prosecution*. Temuan empiris menegaskan bahwa keberadaan bukti elektronik, jumlah alat bukti yang sah, kompleksitas perkara, dan kuantitas saksi merupakan klaster variabel penentu paling kritis dalam memengaruhi hasil peradilan. Penemuan arah pengaruh negatif pada variabel kompleksitas perkara secara langsung memvalidasi teori beban pembuktian (*burden of proof*), membuktikan bahwa peningkatan kerumitan anatomi kasus secara inheren menuntut ambang batas pembuktian yang lebih ketat. Dari aspek praktis dan kebijakan, kerangka kerja ini berpotensi diadopsi sebagai sistem pendukung keputusan (*decision support system*) institusional yang mampu memitigasi risiko kegagalan penuntutan sejak dini serta menekan disparitas penanganan perkara tanpa mendegradasi independensi profesional maupun diskresi jaksa. Meskipun memberikan hasil yang kokoh, penelitian ini memiliki keterbatasan tertentu yang membuka peluang bagi pengembangan di masa depan. Fokus dataset yang berpusat pada karakteristik perkara korupsi berpotensi memerlukan adaptasi arsitektur dan rekayasa fitur ulang jika diterapkan pada domain tindak pidana lainnya. Selain itu, pendekatan interpretasi lokal saat ini masih mengandalkan fitur tabular statis. Oleh karena itu, penelitian selanjutnya disarankan untuk memperluas cakupan model dengan mengintegrasikan teknologi *Natural Language Processing* (NLP) tingkat lanjut guna mengekstraksi informasi semantik langsung dari teks dokumen dakwaan dan berkas perkara yang tidak terstruktur. Pengembangan lebih lanjut juga dapat diarahkan pada pengujian

validitas temporal model secara berkala untuk menjaga akurasi prediksi terhadap dinamika evolusi modus operandi kejahatan keuangan modern.

REFERENSI

- Agrawal, R., Gupta, T., Gupta, S., Chauhan, S., Patel, P., & Hamdare, S. (2025). Fostering Trust and Interpretability: Integrating Explainable AI (XAI) with Machine Learning for Enhanced Disease Prediction and Decision Transparency. *Diagnostic Pathology*, 20(1). <https://doi.org/10.1186/s13000-025-01686-3>
- Ahmed, I., Jeon, G., & Piccialli, F. (2022). From Artificial Intelligence to Explainable Artificial Intelligence in Industry 4.0: A Survey on What, How, and Where. *IEEE Transactions on Industrial Informatics*, 18(8), 5031–5042. <https://doi.org/10.1109/TII.2022.3146552>
- Ali, M., Mulyono, A., & Nurhidayat, S. (2023). The Application of a Human Rights Approach toward Crimes of Corruption: Analyzing Anti-Corruption Regulations and Judicial Decisions. *Laws*, 12(4). <https://doi.org/10.3390/laws12040068>
- Alijoyo, F. A., Janani, S., Santosh, K., Shweihat, S. N., Alshammry, N., Ramesh, J. V. N., & Baker El-Ebiary, Y. A. (2024). Enhancing AI Interpretation and Decision-Making: Integrating Cognitive Computational Models with Deep Learning for Advanced Uncertain Reasoning Systems. *Alexandria Engineering Journal*, 99, 17–30. <https://doi.org/10.1016/j.aej.2024.04.073>
- Androutsopoulou, M., Carayannis, E. G., Askounis, D., & Zotas, N. (2025). Towards AI-Enabled Cyber-Physical Infrastructures, Challenges, Opportunities, and Implications for a Data-Driven eGovernment Theory, Policy, and Practice. *Journal of the Knowledge Economy*. <https://doi.org/10.1007/s13132-025-02726-5>
- Awotunde, J. B., Imoize, A. L., Adeniyi, A. E., Abiodun, K. M., Ayo, E. F., Kavitha, K. V. N., Ajamu, G. J., & Ogundokun, R. O. (2023). Explainable Machine Learning (XML) for Multimedia-Based Healthcare Systems: Opportunities, Challenges, Ethical and Future Prospects. *Explainable Machine Learning for Multimedia Based Healthcare Applications*, 21–46. https://doi.org/10.1007/978-3-031-38036-5_2
- Borgesano, F., De Maio, A., Laghi, P., & Musmanno, R. (2025). Artificial Intelligence and Justice: A Systematic Literature Review and Future Research Perspectives on Justice 5.0. *European Journal of Innovation Management*, 28(11), 349–385. <https://doi.org/10.1108/EJIM-01-2025-0117>
- Cui, J., Shen, X., & Wen, S. (2023). A Survey on Legal Judgment Prediction: Datasets, Metrics, Models and Challenges. *IEEE Access*, 11, 102050–102071. <https://doi.org/10.1109/ACCESS.2023.3317083>
- Esparza-Gómez, J. M., Luque-Vega, L. F., Guerrero-Osuna, H. A., Carrasco-Navarro, R., García-Vázquez, F., Mata-Romero, M. E., Olvera-Olvera, C. A., Carlos-Mancilla, M. A., & Solís-Sánchez, L. O. (2023). Long Short-Term Memory Recurrent Neural Network and Extreme Gradient Boosting Algorithms Applied in a Greenhouse's Internal Temperature Prediction. *Applied Sciences (Switzerland)*, 13(22). <https://doi.org/10.3390/app132212341>
- Fitriyanti, L. D., & Suwandono, A. (2025). Perampasan Aset Sebagai Sanksi Tambahan: Analisis Pengembalian Kerugian Negara dalam Penanganan Tindak Pidana Korupsi di Indonesia. *Jaksa: Jurnal Kajian Ilmu Hukum Dan Politik*, 3(3), 13–27. <https://doi.org/10.51903/N9VPM692>

- Fridkin, S., & Bendersky, M. (2026). Interpretable Machine Learning: A Comprehensive Review of Foundations, Methods, and the Path Forward. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 16(1). <https://doi.org/10.1002/widm.70075>
- Gharagoz, M. M., Noureldin, M., & Kim, J. (2025). Explainable Machine Learning (XML) Framework for Seismic Assessment of Structures using Extreme Gradient Boosting (XGBoost). *Engineering Structures*, 327. <https://doi.org/10.1016/j.engstruct.2025.119621>
- Guido, R., Ferrisi, S., Lofaro, D., & Conforti, D. (2024). An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review. *Information (Switzerland)*, 15(4). <https://doi.org/10.3390/info15040235>
- Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2024). Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. *Cognitive Computation*, 16(1), 45–74. <https://doi.org/10.1007/s12559-023-10179-8>
- Kabir, M. F., Chen, T., & Ludwig, S. A. (2023). A Performance Analysis of Dimensionality Reduction Algorithms in Machine Learning Models for Cancer Prediction. *Healthcare Analytics*, 3. <https://doi.org/10.1016/j.health.2022.100125>
- Khairullah, G., & Hanif, G. (2023). Harmonisasi antara Kepolisian, Kejaksaan dan Komisi Pemberantasan Tindak Pidana Korupsi (KPK) dalam Penyidikan Tindak Pidana Korupsi di X. Perkara: Jurnal Ilmu Hukum dan Politik, 1(2), 14–24. <https://doi.org/10.51903/PERKARA.V1I2.1179>
- Kossay, M. (2026). Jaringan Aktor dalam Penanganan Perkara Korupsi oleh Kejaksaan: Pendekatan Social Network Analysis. Perkara: Jurnal Ilmu Hukum Dan Politik, 4(1), 143–157. <https://doi.org/10.51903/5BK5K553>
- Kyriazos, T., & Poga, M. (2024). Application of Machine Learning Models in Social Sciences: Managing Nonlinear Relationships. *Encyclopedia*, 4(4), 1790–1805. <https://doi.org/10.3390/encyclopedia4040118>
- Laxman, V., Ramesh, N., Jaya Prakash, S. K., & Aluvala, R. (2024). Emerging Threats in Digital Payment and Financial Crime: A Bibliometric Review. *Journal of Digital Economy*, 3, 205–222. <https://doi.org/10.1016/j.jdec.2025.04.002>
- Lee, J., & Dai, Y. (2025). Artificial Intelligence and Data-Driven Technology in the Public Policy Cycle: Comparative Applications, Opportunities, and Risks. *Journal of Comparative Policy Analysis: Research and Practice*, 27(5–6), 477–493. <https://doi.org/10.1080/13876988.2025.2598371>
- Odilla, F. (2023). Bots Against Corruption: Exploring the Benefits and Limitations of AI-Based Anti-Corruption Technology. *Crime, Law and Social Change*, 80(4), 353–396. <https://doi.org/10.1007/s10611-023-10091-0>
- Ortigossa, E. S., Goncalves, T., & Nonato, L. G. (2024). EXplainable Artificial Intelligence (XAI) - From Theory to Methods and Applications. *IEEE Access*, 12, 80799–80846. <https://doi.org/10.1109/ACCESS.2024.3409843>
- Pelosi, D., Cacciagrano, D., & Piangerelli, M. (2025). Explainability and Interpretability in Concept

- and Data Drift: A Systematic Literature Review. *Algorithms*, 18(7). <https://doi.org/10.3390/a18070443>
- Saheb, T., & Saheb, T. (2022). Topical Review of Artificial Intelligence National Policies: A Mixed Method Analysis. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4208097>
- Saranya, A., & Subhashini, R. (2023). A Systematic Review of Explainable Artificial Intelligence Models and Applications: Recent Developments and Future Trends. *Decision Analytics Journal*, 7. <https://doi.org/10.1016/j.dajour.2023.100230>
- Shafighfard, T., Kazemi, F., Bagherzadeh, F., Mieloszyk, M., & Yoo, D. Y. (2024). Chained Machine Learning Model for Predicting Load Capacity and Ductility of Steel Fiber, Reinforced Concrete Beams. *Computer-Aided Civil and Infrastructure Engineering*, 39(23), 3573–3594. <https://doi.org/10.1111/mice.13164>
- Strielkowski, W., Vlasov, A., Selivanov, K., Muraviev, K., & Shakhnov, V. (2023). Prospects and Challenges of the Machine Learning and Data-Driven Methods for the Predictive Analysis of Power Systems: A Review. *Energies*, 16(10). <https://doi.org/10.3390/en16104025>
- Susnjak, T. (2024). Beyond Predictive Learning Analytics Modelling and onto Explainable Artificial Intelligence with Prescriptive Analytics and ChatGPT. *International Journal of Artificial Intelligence in Education*, 34(2), 452–482. <https://doi.org/10.1007/s40593-023-00336-3>
- Theodorou, G., Karagiorgou, S., Modic, J., & De Marco, G. (2025). Transparent Fight Against Corruption Leveraging Advanced and Explainable Data Analytics. *EEITE 2025 - 6th International Conference in Electronic Engineering and Information Technology*. <https://doi.org/10.1109/EEITE65381.2025.11166031>
- Tursunalieva, A., Alexander, D. L. J., Dunne, R., Li, J., Riera, L., & Zhao, Y. (2024). Making Sense of Machine Learning: A Review of Interpretation Techniques and Their Applications. *Applied Sciences (Switzerland)*, 14(2). <https://doi.org/10.3390/app14020496>
- Virkar, S., Alexopoulos, C., Tsekeridou, S., & Novak, A. S. (2022). A User-Centred Analysis of Decision Support Requirements in Legal Informatics. *Government Information Quarterly*, 39(3). <https://doi.org/10.1016/j.giq.2022.101713>
- Wronka, C. (2023). Financial Crime in the Decentralized Finance Ecosystem: New Challenges for Compliance. *Journal of Financial Crime*, 30(1), 97–113. <https://doi.org/10.1108/JFC-09-2021-0218>
- Yadav, D. K., Kaushik, A., & Yadav, N. (2024). Predicting Machine Failures using Machine Learning and Deep Learning Algorithms. *Sustainable Manufacturing and Service Economics*, 3, 100029. <https://doi.org/10.1016/j.smse.2024.100029>
- Yulianingsih, S. (2026). Analysis of the Role of Public Prosecutors in Proving Corruption Crimes in Indonesia. *Jaksa: Jurnal Kajian Ilmu Hukum Dan Politik*, 4(2), 197–210. <https://doi.org/10.51903/4KYE0P36>
- Yunus, A., & Kalamiah, Moh. J. (2023). Pertanggung Jawaban Pidana dalam Tindak Pidana Korupsi yang Dilakukan oleh Hakim. *Hakim: Jurnal Ilmu Hukum Dan Sosial*, 1(4), 341–355. <https://doi.org/10.51903/HAKIM.V1I4.1711>