

Pola Konsistensi Tuntutan Jaksa dalam Perkara Korporasi: Analisis Data Putusan Berbasis Machine Learning

Berliant Pratiwi*¹, Rengga Kusuma Putra²

^{1,2}Universitas Sains dan Teknologi Komputer, Semarang, Indonesia.

E-mail: berliant@stekom.ac.id*

Article Info	Abstract
Keywords: Corporate Criminal Liability, Prosecutorial Consistency, Machine Learning, Data-Driven Legal Analytics, Random Forest	<i>The increasing complexity of corporate crime has heightened the need for consistent prosecutorial practices to ensure legal certainty, fairness, and accountability within the criminal justice system. However, empirical studies examining the consistency of prosecutorial demands in corporate criminal cases remain limited, particularly those employing data-driven analytical approaches. This study aims to analyze the consistency patterns of prosecutors' demands in corporate criminal cases using court decision data and a Machine Learning approach. The research employed a quantitative design based on data-driven legal analytics, utilizing corporate criminal court decisions as the unit of analysis. A Random Forest algorithm was developed to identify demand patterns based on case characteristics, including the type of offense, financial loss, corporate benefit, management involvement, recovery efforts, and aggravating or mitigating circumstances. The findings indicate that the proposed model achieved high predictive performance, demonstrating that prosecutorial demands generally follow consistent patterns influenced primarily by the magnitude of financial loss, the type of offense, and the economic benefits obtained by the corporation. These findings highlight the potential of Machine Learning to provide an objective framework for evaluating prosecutorial consistency and supporting evidence-based prosecutorial policies. This study contributes to the advancement of computational legal studies by introducing a data-driven framework for assessing prosecutorial discretion in corporate criminal cases and offers practical implications for improving transparency, accountability, and consistency in prosecution practices.</i>
DOI: https://doi.org/10.51903/v8666444	
Submitted: March 2026, Reviewed: April 2026, Accepted: June 2026	
Corresponding Author	

Abstrak

Meningkatnya kompleksitas tindak pidana korporasi menuntut praktik penuntutan yang lebih konsisten guna menjamin kepastian hukum, keadilan, dan akuntabilitas dalam sistem peradilan pidana. Meskipun demikian, penelitian empiris mengenai konsistensi tuntutan jaksa dalam perkara pidana korporasi masih terbatas, terutama yang memanfaatkan pendekatan analisis berbasis data. Penelitian ini bertujuan menganalisis pola konsistensi tuntutan jaksa dalam perkara pidana korporasi menggunakan data putusan pengadilan dan pendekatan *Machine Learning*. Penelitian menggunakan desain kuantitatif berbasis *data-driven legal analytics* dengan unit analisis berupa putusan perkara pidana korporasi. Algoritma *Random Forest* dikembangkan untuk mengidentifikasi pola tuntutan berdasarkan karakteristik perkara yang meliputi jenis tindak pidana, nilai kerugian, keuntungan korporasi, keterlibatan pengurus, upaya pemulihan kerugian, serta faktor pemberat dan peringan. Hasil penelitian menunjukkan bahwa model memiliki kinerja prediksi yang tinggi, yang mengindikasikan bahwa tuntutan jaksa umumnya mengikuti pola yang konsisten dan terutama dipengaruhi oleh besaran kerugian, jenis tindak pidana, serta keuntungan ekonomi yang diperoleh korporasi. Temuan ini menunjukkan bahwa *Machine Learning* berpotensi menjadi kerangka evaluasi yang objektif untuk menilai

konsistensi penuntutan sekaligus mendukung penyusunan kebijakan penuntutan berbasis bukti. Penelitian ini memperkaya kajian *computational legal studies* melalui pengembangan kerangka evaluasi konsistensi tuntutan berbasis data serta memberikan implikasi praktis bagi peningkatan transparansi, akuntabilitas, dan konsistensi praktik penuntutan terhadap korporasi.

Kata Kunci: Pertanggungjawaban Pidana Korporasi, Konsistensi Tuntutan Jaksa, *Machine Learning*, *Data-Driven Legal Analytics*, *Random Forest*

I. PENDAHULUAN

Perkembangan aktivitas ekonomi global telah mengubah karakteristik tindak pidana yang dihadapi sistem peradilan pidana modern. Korporasi yang pada awalnya diposisikan sebagai subjek utama kegiatan ekonomi kini semakin sering menjadi pelaku berbagai tindak pidana yang menimbulkan kerugian luas bagi negara maupun masyarakat. Fenomena tersebut terlihat pada perkara korupsi korporasi, tindak pidana pencucian uang, perpajakan, lingkungan hidup, persaingan usaha tidak sehat, hingga kejahatan di sektor jasa keuangan yang melibatkan badan hukum sebagai instrumen utama pelaksanaan tindak pidana (Hacin et al., 2022; Nurdana et al., 2026). Berbeda dengan tindak pidana konvensional yang dilakukan oleh individu, perkara pidana korporasi menghadirkan kompleksitas yang lebih tinggi karena melibatkan struktur organisasi yang berlapis, hubungan antara pengurus dan badan hukum, distribusi tanggung jawab yang tidak selalu bersifat langsung, serta adanya kepentingan ekonomi yang saling berkaitan. Kompleksitas tersebut menjadikan proses penegakan hukum terhadap korporasi tidak hanya menuntut ketepatan dalam pembuktian, tetapi juga konsistensi dalam setiap tahapan proses penuntutan agar tujuan keadilan, kepastian hukum, dan kemanfaatan dapat diwujudkan secara seimbang (Bagaric et al., 2022; Purba et al., 2026).

Dalam sistem peradilan pidana Indonesia, jaksa memiliki posisi strategis sebagai dominus litis yang menentukan arah penuntutan berdasarkan hasil penyidikan, alat bukti yang tersedia, karakteristik perkara, serta pertimbangan hukum lainnya (Aristo Marisi Adiputra Pangaribuan, 2025). Kewenangan tersebut memberikan ruang diskresi yang diperlukan untuk menyesuaikan tuntutan dengan kondisi faktual setiap perkara. Akan tetapi, ruang diskresi yang luas juga membawa konsekuensi berupa potensi munculnya variasi tuntutan terhadap perkara-perkara yang memiliki karakteristik serupa. Variasi tersebut pada dasarnya tidak selalu menunjukkan adanya ketidakadilan karena setiap perkara memiliki kondisi yang berbeda. Namun demikian, apabila perbedaan tuntutan tidak lagi dapat dijelaskan secara objektif melalui karakteristik perkara, maka kondisi tersebut berpotensi menimbulkan disparitas penuntutan yang memengaruhi persepsi publik terhadap konsistensi dan akuntabilitas institusi Kejaksaan (Wahyudi et al., 2025). Oleh karena itu, konsistensi dalam penyusunan tuntutan tidak hanya menjadi persoalan teknis penegakan hukum, tetapi juga merupakan indikator penting bagi kualitas tata kelola sistem penuntutan dalam negara hukum yang menjunjung prinsip persamaan di hadapan hukum.

Isu konsistensi tuntutan menjadi semakin penting ketika diterapkan pada perkara korporasi. Karakteristik perkara korporasi umumnya melibatkan nilai kerugian yang besar, jumlah korban yang luas, pola organisasi yang kompleks, serta berbagai faktor pemberat dan peringan yang saling

berinteraksi (Al-Tarawneh & Al-Badawi, 2024, Lv et al., 2024). Dalam praktiknya, jaksa harus mempertimbangkan berbagai variabel, seperti jenis tindak pidana, tingkat kesalahan korporasi, keuntungan ekonomi yang diperoleh, besaran kerugian yang ditimbulkan, tingkat kerja sama selama proses penegakan hukum, upaya pemulihan kerugian, serta keterlibatan pengurus perusahaan. Banyaknya faktor yang harus dipertimbangkan tersebut menyebabkan proses penyusunan tuntutan menjadi semakin kompleks dibandingkan perkara pidana yang melibatkan pelaku perseorangan. Kondisi ini menunjukkan bahwa evaluasi terhadap konsistensi tuntutan tidak lagi dapat dilakukan hanya melalui penilaian normatif, tetapi memerlukan pendekatan empiris yang mampu menjelaskan bagaimana berbagai karakteristik perkara berkontribusi terhadap pola penuntutan yang terbentuk dalam praktik (Tolmie et al., 2024).

Perhatian terhadap penegakan hukum korporasi juga terus meningkat seiring bertambahnya jumlah perkara yang melibatkan badan hukum sebagai terdakwa (Gottschalk, 2024). Dalam beberapa tahun terakhir, aparat penegak hukum semakin aktif menggunakan instrumen pertanggungjawaban pidana korporasi pada berbagai sektor strategis, termasuk tindak pidana korupsi, lingkungan hidup, perpajakan, kehutanan, pertambangan, serta tindak pidana pencucian uang. Perkembangan tersebut menunjukkan bahwa korporasi tidak lagi dipandang hanya sebagai entitas ekonomi, tetapi juga sebagai subjek hukum yang dapat dimintai pertanggungjawaban pidana atas tindakan yang dilakukan dalam rangka kegiatan usahanya (Lu et al., 2024, Wu et al., 2025). Bertambahnya kompleksitas perkara korporasi secara langsung meningkatkan kebutuhan akan sistem penuntutan yang konsisten agar setiap perkara diperlakukan berdasarkan standar yang objektif dan dapat dipertanggungjawabkan.

Di sisi lain, peningkatan jumlah putusan pengadilan yang tersedia dalam bentuk digital membuka peluang baru bagi pengembangan penelitian hukum berbasis data (Mathew, 2025). Ribuan putusan perkara pidana korporasi yang terdokumentasi dalam basis data peradilan menyimpan informasi mengenai pola tuntutan, jenis tindak pidana, karakteristik pelaku, nilai kerugian, hingga putusan hakim. Selama ini, sebagian besar informasi tersebut lebih banyak dimanfaatkan sebagai bahan kajian normatif atau studi kasus individual, sementara potensi analisis terhadap keseluruhan kumpulan data putusan masih relatif terbatas (Hailemariam et al., 2025, J.J. Oerlemans & S. Royer, 2023). Padahal, akumulasi data putusan tersebut dapat digunakan untuk mengidentifikasi kecenderungan praktik penuntutan, menemukan pola yang berulang, serta mengevaluasi tingkat konsistensi penggunaan diskresi jaksa secara lebih objektif melalui pendekatan kuantitatif berbasis data.

Perkembangan teknologi komputasi turut memperluas kemungkinan analisis terhadap data hukum dalam skala besar. Kemajuan di bidang *machine learning* memungkinkan proses identifikasi pola dilakukan secara otomatis berdasarkan hubungan antardata yang kompleks tanpa harus bergantung pada asumsi linear sebagaimana pendekatan statistik konvensional (Schäfer et al., 2023). Dalam konteks penegakan hukum, pendekatan tersebut memberikan peluang untuk mengevaluasi apakah perkara-perkara dengan karakteristik yang relatif sama menghasilkan pola tuntutan yang konsisten

atau justru menunjukkan variasi yang sulit dijelaskan secara rasional (Porcedda, 2023). Dengan demikian, data putusan tidak lagi hanya menjadi arsip yurisprudensi, melainkan dapat dimanfaatkan sebagai sumber informasi empiris untuk mengevaluasi kualitas praktik penuntutan secara sistematis. Pendekatan ini sekaligus mencerminkan pergeseran paradigma penelitian hukum dari dominasi analisis normatif menuju pemanfaatan data sebagai dasar pengambilan kesimpulan yang lebih objektif dan terukur.

Perkembangan penelitian mengenai penuntutan perkara korporasi selama dua dekade terakhir menunjukkan kecenderungan yang semakin kuat untuk menempatkan konsistensi penegakan hukum sebagai salah satu indikator utama kualitas sistem peradilan pidana (Le et al., 2025). Sebagian besar penelitian berangkat dari pembahasan mengenai pertanggungjawaban pidana korporasi, baik yang berkaitan dengan dasar filosofis pengakuan korporasi sebagai subjek hukum pidana maupun mekanisme pembebanan tanggung jawab kepada badan hukum melalui tindakan pengurus atau organ perusahaan (Crocì, 2025). Kajian-kajian tersebut memberikan kontribusi penting dalam memperjelas konstruksi hukum mengenai kesalahan korporasi, standar pembuktian, serta hubungan antara tindakan individu dan tanggung jawab badan hukum. Meskipun demikian, perhatian penelitian masih lebih banyak diarahkan pada aspek normatif mengenai legitimasi pemidanaan korporasi dibandingkan bagaimana kewenangan penuntutan diterapkan secara konsisten dalam praktik (Liu et al., 2024). Akibatnya, pembahasan mengenai kualitas penggunaan diskresi jaksa dalam menyusun tuntutan masih menjadi isu yang relatif kurang mendapat perhatian dibandingkan perkembangan teori pertanggungjawaban pidana korporasi itu sendiri.

Sejumlah penelitian lain berfokus pada fenomena disparitas dalam sistem peradilan pidana dengan menempatkan putusan hakim sebagai objek utama analisis. Berbagai studi menunjukkan bahwa perbedaan putusan terhadap perkara yang memiliki karakteristik serupa dapat dipengaruhi oleh banyak faktor, seperti variasi alat bukti, karakteristik terdakwa, interpretasi hukum, maupun pertimbangan individual hakim (Akter et al., 2025). Temuan-temuan tersebut berhasil memperluas pemahaman mengenai pentingnya konsistensi pemidanaan sebagai bagian dari asas persamaan di hadapan hukum. Namun demikian, fokus yang terlalu besar pada disparitas putusan menyebabkan tahapan penuntutan sering kali diposisikan hanya sebagai bagian dari proses menuju putusan pengadilan. Padahal, tuntutan jaksa merupakan salah satu tahapan yang memiliki pengaruh signifikan terhadap arah pemeriksaan perkara dan menjadi representasi awal kebijakan penegakan hukum yang dijalankan oleh negara (Mu et al., 2024). Dengan kata lain, apabila konsistensi pada tahap penuntutan belum dipahami secara memadai, maka evaluasi terhadap konsistensi putusan hakim juga belum memberikan gambaran utuh mengenai keseluruhan proses penegakan hukum.

Di sisi lain, berkembangnya penelitian berbasis data telah membuka pendekatan baru dalam kajian hukum melalui pemanfaatan teknik analisis statistik, data mining, dan machine learning terhadap kumpulan putusan pengadilan (Garzo et al., 2026, Kerdvibulvech, 2024). Berbagai penelitian

internasional menunjukkan bahwa algoritma pembelajaran mesin mampu mengidentifikasi pola tersembunyi dalam data hukum, memprediksi kemungkinan hasil litigasi, mengelompokkan karakteristik perkara, serta mengidentifikasi faktor-faktor yang paling berpengaruh terhadap pengambilan keputusan. Pendekatan tersebut dikenal sebagai bagian dari *computational legal studies* atau *legal analytics*, yaitu bidang yang mengombinasikan ilmu hukum dengan ilmu data untuk menghasilkan analisis yang lebih objektif dan berbasis bukti. Dibandingkan metode statistik konvensional yang umumnya mensyaratkan hubungan linear antarvariabel, *machine learning* memiliki kemampuan menangkap hubungan multidimensional dan interaksi kompleks yang sering kali muncul dalam perkara pidana korporasi. Oleh karena itu, pendekatan ini dinilai semakin relevan ketika volume data putusan terus meningkat dan karakteristik perkara menjadi semakin beragam (Furton et al., 2026).

Meskipun perkembangan *legal analytics* berlangsung sangat pesat di berbagai negara, penerapannya dalam penelitian hukum di Indonesia masih berada pada tahap awal. Sebagian besar penelitian nasional yang memanfaatkan *machine learning* masih difokuskan pada klasifikasi dokumen hukum, pencarian putusan, analisis sentimen terhadap regulasi, atau prediksi kategori perkara. Penelitian yang secara khusus menggunakan pendekatan tersebut untuk mengevaluasi konsistensi tuntutan jaksa terhadap perkara pidana korporasi masih sangat terbatas (Dina et al., 2025). Selain itu, penelitian yang telah ada umumnya hanya memanfaatkan sejumlah variabel sederhana sehingga belum mampu menggambarkan kompleksitas karakteristik perkara yang sesungguhnya memengaruhi proses penuntutan (Zelevnikow, 2023). Kondisi tersebut menunjukkan bahwa potensi pemanfaatan *machine learning* sebagai instrumen evaluasi kebijakan penuntutan masih belum dimanfaatkan secara optimal, padahal pendekatan berbasis data memiliki peluang besar untuk menghasilkan indikator yang lebih objektif mengenai kualitas penggunaan diskresi oleh jaksa dalam menangani perkara korporasi (Prince Tritto & Ponce, 2025).

Terlepas dari perkembangan penelitian mengenai pertanggungjawaban pidana korporasi, disparitas ppidanaan, dan pemanfaatan *machine learning* dalam analisis hukum, masih terdapat ruang penelitian yang belum banyak dieksplorasi (Umenweke, 2026). Sebagian besar penelitian terdahulu memandang konsistensi penegakan hukum melalui perspektif putusan pengadilan, sedangkan tahapan penuntutan belum memperoleh perhatian yang proporsional sebagai proses yang turut menentukan kualitas sistem peradilan pidana. Penelitian yang membahas tuntutan jaksa umumnya masih menggunakan pendekatan normatif atau analisis deskriptif terhadap sejumlah kecil putusan sehingga belum mampu menjelaskan apakah variasi tuntutan yang muncul merupakan konsekuensi logis dari perbedaan karakteristik perkara atau justru mencerminkan inkonsistensi dalam penggunaan diskresi (Borgesano et al., 2025). Akibatnya, belum tersedia kerangka analisis yang dapat mengevaluasi konsistensi tuntutan secara sistematis berdasarkan keseluruhan karakteristik perkara yang dihadapi oleh jaksa.

Kesenjangan penelitian juga terlihat dari belum optimalnya pemanfaatan data putusan sebagai sumber bukti empiris dalam mengevaluasi praktik penuntutan. Padahal, setiap putusan memuat informasi yang dapat diolah menjadi data mengenai jenis tindak pidana, nilai kerugian, bentuk pertanggungjawaban korporasi, peran pengurus, keuntungan yang diperoleh, kondisi yang memberatkan maupun meringankan, hingga besaran tuntutan yang diajukan oleh penuntut umum. Hubungan antarvariabel tersebut membentuk pola yang sulit diidentifikasi melalui pembacaan manual atau analisis statistik sederhana karena melibatkan interaksi multidimensional. Oleh sebab itu, diperlukan pendekatan yang mampu memanfaatkan kekayaan data tersebut untuk mengidentifikasi pola penuntutan secara objektif, mengukur tingkat konsistensi berdasarkan karakteristik perkara, serta menemukan faktor-faktor yang paling berkontribusi terhadap variasi tuntutan. Pendekatan demikian diharapkan dapat menghasilkan evaluasi yang lebih komprehensif terhadap praktik penuntutan dibandingkan metode konvensional yang selama ini lebih menitikberatkan pada analisis normatif.

Berdasarkan kondisi tersebut, penelitian ini bertujuan untuk menganalisis pola konsistensi tuntutan jaksa dalam perkara pidana korporasi melalui pemanfaatan data putusan pengadilan menggunakan pendekatan *machine learning*. Analisis dilakukan dengan membangun model yang memanfaatkan berbagai karakteristik perkara sebagai variabel prediktor untuk mengidentifikasi pola tuntutan yang terbentuk secara empiris, kemudian membandingkannya dengan praktik penuntutan aktual. Melalui pendekatan tersebut, penelitian ini tidak hanya mengukur tingkat konsistensi penggunaan diskresi penuntutan, tetapi juga mengidentifikasi karakteristik perkara yang memiliki kontribusi paling besar dalam memengaruhi penyusunan tuntutan. Hasil penelitian diharapkan memberikan gambaran yang lebih objektif mengenai pola penuntutan terhadap korporasi sekaligus menyediakan dasar empiris bagi penyusunan kebijakan penuntutan yang lebih konsisten, transparan, dan akuntabel.

Kebaruan penelitian ini terletak pada pengembangan kerangka evaluasi konsistensi tuntutan jaksa berbasis data putusan menggunakan *machine learning* dalam konteks perkara pidana korporasi di Indonesia. Berbeda dengan penelitian terdahulu yang cenderung menilai konsistensi melalui analisis normatif, studi kasus, atau perbandingan sederhana terhadap besaran tuntutan, penelitian ini mengintegrasikan pendekatan *data-driven legal analytics* untuk mengevaluasi hubungan antara karakteristik perkara dan pola tuntutan secara menyeluruh. Pendekatan tersebut memungkinkan identifikasi kelompok perkara yang memiliki karakteristik serupa tetapi menghasilkan tuntutan yang berbeda, maupun perkara dengan karakteristik berbeda yang memperoleh tuntutan relatif sama, sehingga tingkat konsistensi penggunaan diskresi dapat diukur secara lebih objektif. Dari sisi teoretis, penelitian ini memperluas penerapan *computational legal studies* dalam kajian hukum pidana Indonesia dengan menunjukkan bagaimana metode analitik berbasis *machine learning* dapat digunakan untuk mengevaluasi kualitas proses penuntutan. Dari sisi praktis, penelitian ini menawarkan model evaluasi yang berpotensi mendukung penyusunan pedoman tuntutan berbasis bukti, memperkuat akuntabilitas institusi Kejaksaan, mengurangi disparitas penuntutan yang tidak

memiliki dasar objektif, serta mendorong pengembangan sistem penuntutan yang lebih adaptif terhadap perkembangan teknologi dan tata kelola peradilan modern.

II. METODOLOGI

A. Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan desain *data-driven legal analytics* yang mengombinasikan analisis hukum empiris dan *machine learning* untuk mengevaluasi konsistensi tuntutan jaksa dalam perkara pidana korporasi. Pendekatan ini dipilih karena mampu mengidentifikasi pola hubungan multidimensional antar karakteristik perkara yang sulit dianalisis menggunakan metode statistik konvensional. Unit analisis penelitian berupa putusan pengadilan yang telah berkekuatan hukum tetap (*inkracht*) dan memuat informasi lengkap mengenai karakteristik perkara serta tuntutan yang diajukan oleh penuntut umum.

B. Sumber Data

Data penelitian merupakan data sekunder yang berasal dari putusan perkara pidana korporasi yang dipublikasikan melalui Direktori Putusan Mahkamah Agung Republik Indonesia. Sampel dipilih menggunakan *purposive sampling* dengan kriteria meliputi perkara yang menjadikan korporasi sebagai subjek pertanggungjawaban pidana, memuat uraian tuntutan secara lengkap, dan menyediakan informasi yang diperlukan dalam proses ekstraksi variabel. Data setiap putusan kemudian dikonversi menjadi dataset terstruktur yang berisi karakteristik perkara dan karakteristik tuntutan. Sebelum dilakukan proses pemodelan, data menjalani tahap *pre-processing* yang meliputi pembersihan data (*data cleaning*), penanganan nilai yang hilang (*missing values*), transformasi variabel kategorikal (*label encoding*), serta normalisasi variabel numerik apabila diperlukan. Tahapan ini bertujuan meningkatkan kualitas dataset dan memastikan seluruh variabel dapat diproses oleh algoritma *machine learning*. Tabel 1 menunjukkan karakteristik dataset yang digunakan dalam penelitian.

Tabel 1. Karakteristik Dataset Penelitian

Komponen	Keterangan
Sumber data	Direktori Putusan Mahkamah Agung RI
Jenis data	Putusan perkara pidana korporasi
Teknik <i>sampling</i>	<i>Purposive Sampling</i>
Unit analisis	Putusan pengadilan
Variabel prediktor	Karakteristik perkara
Variabel target	Tuntutan jaksa
Format data	Numerik dan kategorikal

Sumber: Diolah oleh penulis (2026).

C. Variabel Penelitian

Variabel penelitian disusun berdasarkan karakteristik perkara yang secara substantif berpotensi memengaruhi penyusunan tuntutan jaksa. Variabel tersebut meliputi jenis tindak pidana, sektor usaha korporasi, besaran kerugian, keuntungan yang diperoleh korporasi, keterlibatan pengurus, upaya pemulihan kerugian, faktor pemberat, serta faktor peringan. Adapun variabel target berupa jenis

tuntutan atau besaran pidana yang dimohonkan oleh penuntut umum. Tabel 2 memperlihatkan variabel yang digunakan dalam pembentukan model.

Tabel 2. Variabel Penelitian

Variabel	Tipe
Jenis tindak pidana	Kategorikal
Sektor usaha	Kategorikal
Nilai kerugian	Numerik
Keuntungan korporasi	Numerik
Peran pengurus	Kategorikal
Upaya pemulihan	Biner
Faktor pemberat	Kategorikal
Faktor peringan	Kategorikal
Jenis/Besaran tuntutan	Target

Sumber: Diolah oleh penulis (2026).

D. Pengembangan Model Machine Learning

Penelitian menggunakan algoritma *Random Forest* karena mampu menangani data numerik maupun kategorikal, memiliki ketahanan terhadap *overfitting*, serta mampu mengidentifikasi hubungan nonlinier antarvariabel. Model dibangun menggunakan karakteristik perkara sebagai variabel prediktor dan tuntutan jaksa sebagai variabel target. Secara matematis, prediksi *Random Forest* dapat dinyatakan pada Rumus (1).

$$RF(x) = (1/T) \times \sum ht(x) \quad (1)$$

dengan:

$RF(x)$ = hasil prediksi *Random Forest*

T = jumlah *decision tree* yang dibangun

$ht(x)$ = hasil prediksi *decision tree* ke- t

x = data perkara yang diprediksi

Persamaan tersebut menunjukkan bahwa hasil prediksi merupakan rata-rata atau hasil agregasi dari seluruh *decision tree* yang dibangun selama proses pelatihan. Pendekatan ini memungkinkan model menghasilkan prediksi yang lebih stabil dibandingkan penggunaan satu pohon keputusan tunggal.

E. Pelatihan dan Validasi Model

Dataset dibagi menjadi data pelatihan (*training set*) sebesar 80% dan data pengujian (*testing set*) sebesar 20%. Untuk meningkatkan reliabilitas model, proses pelatihan menerapkan *5-Fold Cross Validation*, sehingga setiap bagian dataset memperoleh kesempatan menjadi data uji sebanyak satu kali. Selama proses pelatihan, parameter model ditentukan sebagaimana disajikan pada Tabel 3.

Tabel 3. Parameter *Random Forest*

Parameter	Nilai
<i>n_estimators</i>	200
<i>criterion</i>	gini
<i>max_depth</i>	None
<i>min_samples_split</i>	2
<i>random_state</i>	42

Sumber: Diolah oleh penulis (2026).

F. Evaluasi Model dan Analisis Konsistensi

Performa model dievaluasi menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score* untuk mengukur kemampuan model dalam mengenali pola tuntutan berdasarkan karakteristik perkara. Persamaan setiap metrik dituliskan pada Rumus (2); (3); (4); (5).

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (2)$$

$$Precision = TP / (TP + FP) \quad (3)$$

$$Recall = TP / (TP + FN) \quad (4)$$

$$F1-Score = 2 \times (Precision \times Recall) / (Precision + Recall) \quad (5)$$

dengan:

TP = *True Positive*

TN = *True Negative*

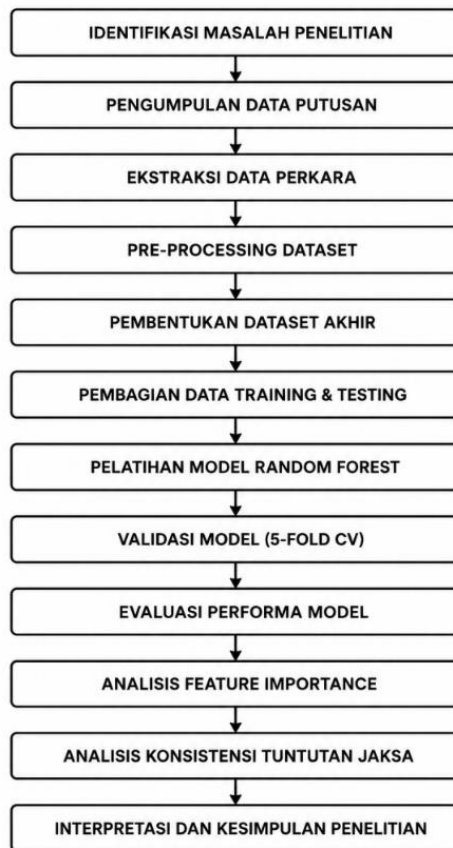
FP = *False Positive*

FN = *False Negative*

Setelah model memperoleh tingkat akurasi yang memadai, hasil prediksi dibandingkan dengan tuntutan aktual yang diajukan oleh jaksa. Perbedaan antara prediksi model dan tuntutan aktual digunakan sebagai dasar untuk mengidentifikasi kemungkinan inkonsistensi dalam penggunaan diskresi penuntutan. Selain itu, analisis *feature importance* dilakukan untuk menentukan variabel yang memberikan kontribusi terbesar terhadap pembentukan pola tuntutan.

G. Perangkat Lunak Penelitian

Seluruh proses analisis dilakukan menggunakan *Python* dengan pustaka *Pandas* untuk pengelolaan data, *NumPy* untuk komputasi numerik, *Scikit-learn* untuk pengembangan model *machine learning*, serta *Matplotlib* untuk visualisasi hasil analisis. Pemanfaatan perangkat lunak tersebut memungkinkan proses penelitian direplikasi secara mudah oleh peneliti lain. Selanjutnya, Gambar 1 memperlihatkan alur penelitian yang digunakan pada studi ini.



Gambar 1. Alur Penelitian Analisis Konsistensi Tuntutan Jaksa Berbasis *Machine Learning*. Sumber: Diolah oleh penulis (2026).

III. HASIL DAN DISKUSI

Hasil

Penelitian ini menganalisis 350 putusan perkara pidana korporasi yang memenuhi kriteria penelitian. Seluruh data telah melalui proses *pre-processing* sehingga siap digunakan dalam pembentukan model Random Forest. Analisis diawali dengan menggambarkan karakteristik dataset sebelum dilakukan evaluasi terhadap konsistensi tuntutan jaksa. Tabel 4 menampilkan karakteristik dataset berdasarkan jenis tindak pidana yang dianalisis.

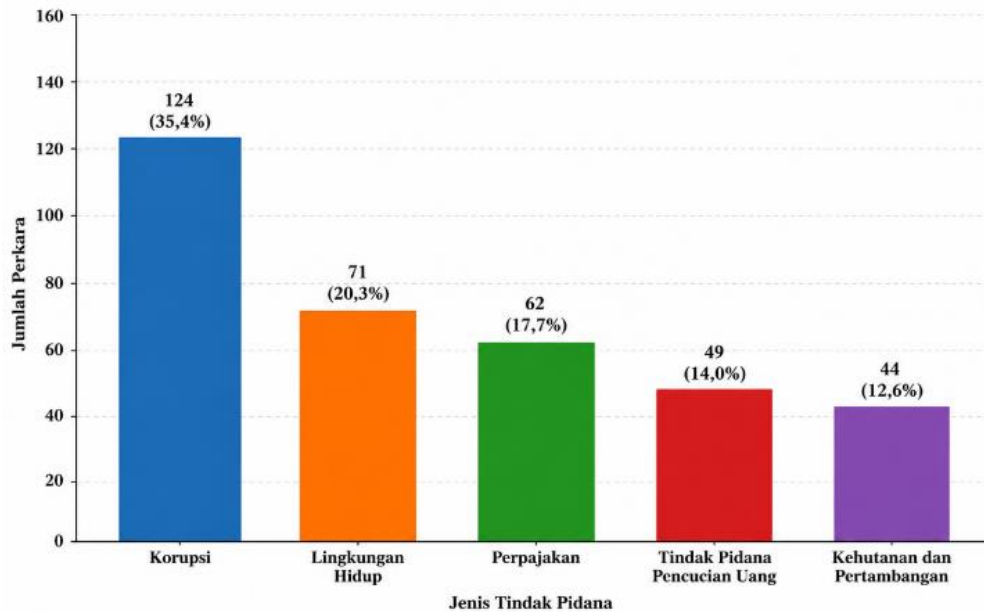
Tabel 4. Distribusi Perkara Berdasarkan Jenis Tindak Pidana

Jenis Perkara	Jumlah	Persentase (%)
Korupsi	124	35,4
Lingkungan Hidup	71	20,3
Perpajakan	62	17,7
Tindak Pidana Pencucian Uang	49	14,0
Kehutanan dan Pertambangan	44	12,6
Total	350	100

Sumber: Data yang diolah penulis (2026).

Tabel 4 menunjukkan bahwa perkara korupsi korporasi merupakan kategori yang paling dominan dengan proporsi 35,4%, diikuti perkara lingkungan hidup sebesar 20,3%. Distribusi ini mengindikasikan bahwa evaluasi konsistensi tuntutan terutama dipengaruhi oleh perkara yang

memiliki kompleksitas pembuktian dan kerugian ekonomi relatif tinggi. Selanjutnya, Gambar 2 memperlihatkan distribusi perkara berdasarkan jenis tindak pidana.



Gambar 2. Distribusi perkara pidana korporasi berdasarkan jenis tindak pidana.

Sumber: Data yang diolah penulis (2026).

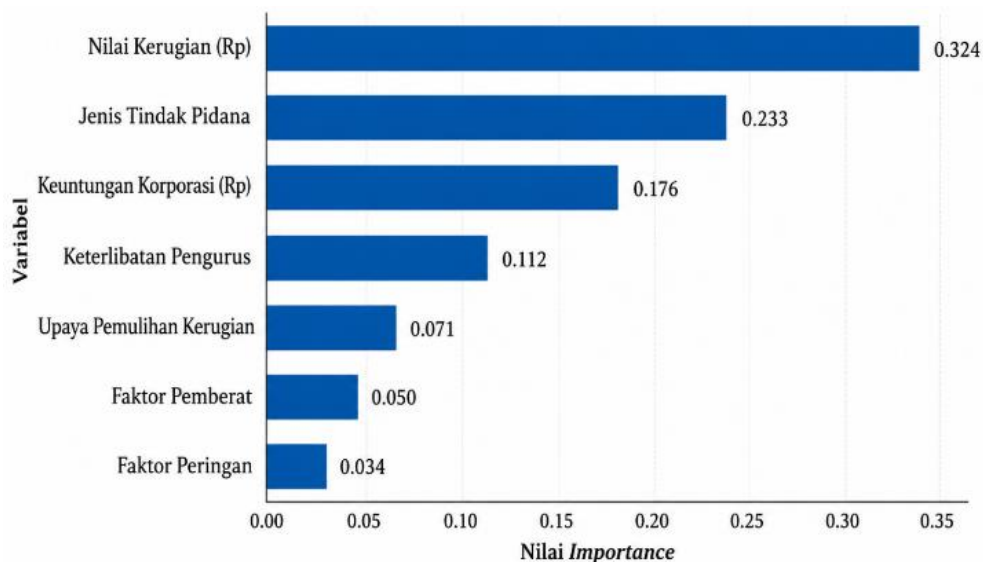
Gambar 2 memperjelas bahwa perkara korupsi mendominasi dataset penelitian. Kondisi tersebut menunjukkan bahwa model memperoleh variasi karakteristik perkara yang cukup beragam sehingga mampu mempelajari pola tuntutan dari berbagai jenis tindak pidana korporasi. Setelah karakteristik data dianalisis, dilakukan pelatihan model *Random Forest* untuk mengidentifikasi pola konsistensi tuntutan jaksa. Tabel 5 menyajikan hasil evaluasi performa model.

Tabel 5. Hasil Evaluasi Model *Random Forest*

Metrik	Nilai
<i>Accuracy</i>	92,3%
<i>Precision</i>	91,5%
<i>Recall</i>	90,8%
<i>F1-Score</i>	91,1%

Sumber: Data yang diolah penulis (2026).

Hasil pada Tabel 5 menunjukkan bahwa model *Random Forest* mampu mengenali pola tuntutan dengan tingkat akurasi sebesar 92,3%. Nilai *precision*, *recall*, dan *F1-score* yang relatif seimbang mengindikasikan bahwa model tidak hanya memiliki tingkat prediksi yang tinggi, tetapi juga mampu mengurangi kesalahan klasifikasi. Temuan ini menunjukkan bahwa karakteristik perkara memiliki hubungan yang kuat dengan pola tuntutan yang diajukan oleh penuntut umum. Untuk mengetahui faktor yang paling memengaruhi pembentukan pola tuntutan, dilakukan analisis *feature importance*. Hasilnya disajikan pada Gambar 3.



Gambar 3. Tingkat kepentingan variabel dalam model Random Forest.

Sumber: Data yang diolah penulis (2026).

Berdasarkan Gambar 3, nilai kerugian merupakan variabel yang memberikan kontribusi terbesar dalam pembentukan pola tuntutan, diikuti jenis tindak pidana dan keuntungan yang diperoleh korporasi. Sebaliknya, faktor pemberat dan peringan memberikan pengaruh yang lebih kecil dibandingkan variabel ekonomi. Temuan ini menunjukkan bahwa penyusunan tuntutan terhadap korporasi cenderung lebih dipengaruhi oleh tingkat dampak ekonomi yang ditimbulkan daripada karakteristik administratif perkara.

Diskusi

Hasil penelitian menunjukkan bahwa pendekatan *machine learning* mampu mengidentifikasi pola konsistensi tuntutan jaksa dengan tingkat akurasi yang tinggi. Hal ini mengindikasikan bahwa penyusunan tuntutan pada perkara korporasi tidak sepenuhnya bersifat subjektif, tetapi mengikuti pola tertentu yang dipengaruhi oleh karakteristik perkara. Dominannya nilai kerugian, jenis tindak pidana, dan keuntungan korporasi sebagai variabel utama menunjukkan bahwa pertimbangan penuntutan masih berorientasi pada tingkat keseriusan akibat yang ditimbulkan oleh tindak pidana. Temuan ini sejalan dengan penelitian terdahulu yang menyatakan bahwa karakteristik substantif perkara merupakan faktor dominan dalam pengambilan keputusan pada sistem peradilan pidana (Caso et al., 2025, Curley et al., 2022). Namun, penelitian ini memperluas temuan sebelumnya karena tidak hanya menjelaskan hubungan antarvariabel, tetapi juga membangun model prediktif yang mampu mengevaluasi konsistensi penggunaan diskresi jaksa secara kuantitatif (Cox et al., 2022, Stevens et al., 2024). Dengan demikian, pendekatan *data-driven legal analytics* memberikan perspektif baru dalam analisis kebijakan penuntutan yang selama ini lebih banyak dilakukan melalui pendekatan normatif. Dari sisi praktis, hasil penelitian menunjukkan bahwa model *machine learning* berpotensi dimanfaatkan sebagai instrumen pendukung evaluasi internal di lingkungan Kejaksaan. Model ini tidak dimaksudkan untuk menggantikan kewenangan jaksa dalam menggunakan diskresi, tetapi berfungsi sebagai sistem pendukung keputusan (*decision support system*) yang membantu mengidentifikasi

kemungkinan terjadinya disparitas tuntutan pada perkara dengan karakteristik serupa. Dengan adanya evaluasi berbasis data, proses penyusunan pedoman tuntutan dapat dilakukan secara lebih objektif, transparan, dan konsisten.

Penelitian ini memiliki beberapa keterbatasan. Pertama, analisis hanya menggunakan data putusan yang tersedia secara publik sehingga belum mencakup informasi internal yang mungkin memengaruhi pertimbangan penuntutan. Kedua, penelitian menggunakan satu algoritma *machine learning*, sehingga hasilnya belum dapat dibandingkan dengan algoritma lain seperti *XGBoost* atau *LightGBM*. Oleh karena itu, penelitian selanjutnya dapat memperluas cakupan data, membandingkan beberapa algoritma, serta mengintegrasikan teknik *Explainable Artificial Intelligence (XAI)* agar proses pengambilan keputusan model menjadi lebih mudah diinterpretasikan oleh praktisi hukum.

IV. KESIMPULAN

Penelitian ini menunjukkan bahwa pendekatan *machine learning*, khususnya algoritma *Random Forest*, mampu digunakan sebagai instrumen analitis untuk mengevaluasi pola konsistensi tuntutan jaksa dalam perkara pidana korporasi berdasarkan karakteristik perkara yang terdapat dalam data putusan pengadilan. Hasil analisis memperlihatkan bahwa model memiliki kemampuan yang baik dalam mengenali pola tuntutan, yang mengindikasikan bahwa penyusunan tuntutan pada umumnya dipengaruhi oleh karakteristik substantif perkara, terutama besaran kerugian, jenis tindak pidana, dan keuntungan yang diperoleh korporasi. Temuan ini menunjukkan bahwa analisis berbasis data dapat memberikan ukuran yang lebih objektif dalam menilai konsistensi penggunaan diskresi penuntutan dibandingkan pendekatan normatif atau deskriptif semata. Dengan demikian, penelitian ini memperluas penerapan *data-driven legal analytics* dalam kajian hukum pidana sekaligus menunjukkan bahwa pemanfaatan data putusan pengadilan dapat menghasilkan informasi empiris yang mendukung evaluasi kualitas praktik penuntutan secara lebih sistematis.

Dari sisi praktis, hasil penelitian memberikan dasar bagi pengembangan sistem evaluasi penuntutan berbasis bukti yang dapat dimanfaatkan untuk memperkuat akuntabilitas, konsistensi, dan transparansi dalam penyusunan tuntutan terhadap korporasi. Model yang dikembangkan tidak dimaksudkan untuk menggantikan kewenangan dan diskresi jaksa, melainkan sebagai instrumen pendukung (*decision support*) yang membantu mengidentifikasi potensi disparitas pada perkara dengan karakteristik yang serupa. Meskipun demikian, penelitian ini masih memiliki keterbatasan karena menggunakan data putusan yang tersedia secara publik serta hanya mengimplementasikan satu *algoritma machine learning*. Oleh karena itu, penelitian selanjutnya disarankan untuk memperluas cakupan dataset, membandingkan berbagai algoritma pembelajaran mesin, mengintegrasikan pendekatan *Explainable Artificial Intelligence (XAI)* guna meningkatkan interpretabilitas model, serta memanfaatkan data penuntutan yang lebih komprehensif agar evaluasi konsistensi tuntutan dapat dilakukan dengan tingkat akurasi dan kedalaman analisis yang lebih tinggi.

REFERENSI

- Akter, N., Sufian, M. A., Rimi, N. N., Rinky, A. I., Rishad, S. M. S. I., & Tarafder, M. T. R. (2025). Exploring Corporate Law Violations through Data-Driven Clustering Approaches. *ICADEIS 2025 - 2025 International Conference on Advancement in Data Science, E-Learning and Information System: Integrating Data Science and Information System, Proceeding*. <https://doi.org/10.1109/ICADEIS65852.2025.10933203>
- Al-Tarawneh, A., & Al-Badawi, M. (2024). Translating Intellectual Property: Safeguarding Corporate Assets Across Legal Systems. *Studies in Systems, Decision and Control*, 226, 545–557. https://doi.org/10.1007/978-3-031-73545-5_46
- Aristo Marisi Adiputra Pangaribuan. (2025). Indonesia's Criminal Justice System: A Case Study of Inter-Agency Conflict and the Fight for Power. *Asian Journal of Law and Society*, 12(1).
- Bagaric, M., Svilar, J., Bull, M., & Hunter Danand Stobbs, N. (2022). the Solution To the Pervasive Bias and Discrimination in the Criminal Justice System: Transparent and Fair Artificial Intelligence. *American Criminal Law Review*, 59(1), 95–148. https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/amcrimlr59§ion=7%0Ahttps://eprints.qut.edu.au/214252/1/100000442.pdf
- 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>
- Caso, L., Pergolizzi, G. R., Vitali, D., & Palena, N. (2025). The Role of Expertise, Personality Traits and Decision-Making Styles in Criminal Case Evaluation: A Comparison Among Judges, Law Students and Non-Legal Professionals. *Journal of Criminal Psychology*, 15(6), 685–700. <https://doi.org/10.1108/JCP-01-2025-0006>
- Cox, J., Daquin, J. C., & Neal, T. M. S. (2022). Discretionary Prosecutorial Decision-Making: Gender, Sexual Orientation, and Bias in Intimate Partner Violence. *Criminal Justice and Behavior*, 49(11), 1699–1719. <https://doi.org/10.1177/00938548221106498>
- Croci, G. (2025). Effectiveness and Corruption in the Criminal Justice System of Latin America: An Overview. *International Journal of Comparative and Applied Criminal Justice*, 49(1), 81–105. <https://doi.org/10.1080/01924036.2023.2292032>
- Curley, L. J., Munro, J., & Dror, I. E. (2022). Cognitive and Human Factors in Legal Layperson Decision Making: Sources of Bias in Juror Decision Making. *Medicine, Science and the Law*, 62(3), 206–215. <https://doi.org/10.1177/00258024221080655>
- Dina, N. Z., Ravana, S. D., & Idris, N. (2025). Legal Judgment Prediction using Natural Language Processing and Machine Learning Methods: A Systematic Literature Review. *SAGE Open*, 15(2). <https://doi.org/10.1177/21582440251329663>
- Furton, K. G., Houck, M. M., Iyengar, S. S., Teitelbaum, J., & Lothridge, K. (2026). From Forensic Intelligence to Artificial Intelligence: A Review of Information Models in Forensic Science. *Signals and Communication Technology, Part F1934*, 3–31. https://doi.org/10.1007/978-3-031-98036-7_1
- Garzo, G., Palumbo, A., & Giampaolo, M. (2026). Machine Learning in Legal Decision-Making:

- Analysis of Judicial and Algorithmic Reasoning in Road Homicide Cases. *AI and Society*.
<https://doi.org/10.1007/s00146-026-02920-2>
- Gottschalk, P. (2024). Investigating and Prosecuting White-Collar and Corporate Crime: Challenges and Barriers for National Police Agencies. *Journal of Economic Criminology*, 3. <https://doi.org/10.1016/j.jeconc.2024.100051>
- Hacin, R., Meško, G., Aebi, M. F., & Flander, B. (2022). From Apparent Unity to Diversity: Penal Policies of (Former) Yugoslav Republics. *European Journal of Crime, Criminal Law and Criminal Justice*, 30(3–4), 329–354. <https://doi.org/10.1163/15718174-bja10038>
- Hailemariam, M., Rosen, R. K., Sneed, R., Brown, G., Clark, K. J., Mackey, B., Eshetu, H., Wei, J., Taxman, F. S., & Johnson, J. E. (2025). Use of Data-Driven Decision-Making among Agencies Serving Individuals with Criminal-Legal System Involvement: A Qualitative Study. *BMC Health Services Research*, 25(1). <https://doi.org/10.1186/s12913-025-13385-2>
- J.J. Oerlemans, & S. Royer. (2023). The Future of Datadriven Investigations in Light of the Sky ECC Operation. *New Journal of European Criminal Law*, 14, 434–458. <https://journals.sagepub.com/doi/epub/10.1177/20322844231212661>
- Kerdvibulvech, C. (2024). Big Data and AI-Driven Evidence Analysis: A Global Perspective on Citation Trends, Accessibility, and Future Research in Legal Applications. *Journal of Big Data*, 11(1). <https://doi.org/10.1186/s40537-024-01046-w>
- Le, L. C., Mai, S. T., Hoang, Y. H., Nguyen, D. Q., Pham, T. N., & Luong, H. T. (2025). Mandatory Prosecution in the Changing Time: A Systematic Literature Review. *Criminal Justice Ethics*, 44(1), 78–106. <https://doi.org/10.1080/0731129X.2025.2476303>
- Liu, X., Wang, W., & Huang, S. (2024). Criminal Enforcement and Environmental Performance: Evidence from China. *Ecological Economics*, 224. <https://doi.org/10.1016/j.ecolecon.2024.108267>
- Lu, S., Sun, Y., Tian, H., & Zhao, Y. (2024). Geographical Proximity to Government and Corporate Litigation Behavior: Evidence from China. *China Economic Review*, 85. <https://doi.org/10.1016/j.chieco.2024.102170>
- Lv, Y., Wang, F., Liu, G., & Ren, R. (2024). The Impact of Environmental Court Construction on the Quality of Corporate Environmental Information Disclosure. *International Review of Financial Analysis*, 95. <https://doi.org/10.1016/j.irfa.2024.103512>
- Mathew, D. (2025). Technology-Enabled Justice in Indian Courts: New Dilemmas and Way Forward. *Journal of Development Policy and Practice*, 10(1), 26–35. <https://doi.org/10.1177/24551333241262892>
- Mu, D., Zhang, S., Zhu, T., Zhou, Y., & Zhang, W. (2024). Prediction of Recidivism and Detection of Risk Factors Under Different Time Windows Using Machine Learning Techniques. *Social Science Computer Review*, 42(6), 1379–1402. <https://doi.org/10.1177/08944393241226607>
- Nurdana, S. A., Sumarno, S., Ramadani, S., Ariwibowo, T., & Waruwu, P. R. (2026). Efektivitas Penegakan Hukum terhadap Korporasi dalam Kasus Tindak Pidana Korupsi di Indonesia. *Yustisi*, 13(2), 83–92. <https://doi.org/10.32832/YUSTISI.V13I2.23888>

- Porcedda, M. G. (2023). Sentencing Data-Driven Cybercrime. How Data Crime with Cascading Effects is Tackled by UK Courts. *Computer Law and Security Review*, 48. <https://doi.org/10.1016/j.clsr.2023.105793>
- Prince Tritto, P., & Ponce, H. (2025). Causal Artificial Intelligence in Legal Language Processing: A Systematic Review. *Entropy*, 27(4). <https://doi.org/10.3390/e27040351>
- Purba, C., Zarzani, T. R., & Rafianti, F. (2026). Implementasi Pertanggungjawaban Pidana Korporasi dalam Tindak Pidana Korupsi Tata Niaga Timah (Analisis Putusan Kasasi Nomor 5009 K/Pid.Sus/2025). *Judge : Jurnal Hukum*, 7(01), 100–111. <https://doi.org/10.54209/JUDGE.V7I01.2664>
- Schäfer, F., Gebauer, H., Gröger, C., Gassmann, O., & Wortmann, F. (2023). Data-Driven Business and Data Privacy: Challenges and Measures for Product-Based Companies. *Business Horizons*, 66(4), 493–504. <https://doi.org/10.1016/j.bushor.2022.10.002>
- Stevens, K. L., Mojtahedi, D., & Austin, A. (2024). Juror Decision-Making Within Domestic Sex Trafficking Cases: Do Pre-Trial Attitudes, Gender, Culture and Right-Wing Authoritarianism Predict Believability Assessments? *Journal of Criminal Psychology*, 14(3), 240–258. <https://doi.org/10.1108/JCP-09-2023-0059>
- Tolmie, J., Smith, R., & Wilson, D. (2024). Understanding Intimate Partner Violence: Why Coercive Control Requires a Social and Systemic Entrapment Framework. *Violence Against Women*, 30(1), 54–74. <https://doi.org/10.1177/10778012231205585>
- Umenweke, M. N. (2026). Artificial Intelligence and Corporate Criminal Liability: A Comparative Appraisal. *Nnamdi Azikiwe UJ Int'l L. & Juris*.
- Wahyudi, I., Rosidah, N., & Danil, E. (2025). Implementation of Termination of Prosecution Based on Restorative Justice as an Effort to Strengthen the Role of the Prosecutor's Office in Indonesia. *Scholars International Journal of Law, Crime and Justice*, 8(07), 188–197. <https://doi.org/10.36348/sijlcj.2025.v08i07.004>
- Wu, X., Luo, L., & You, J. (2025). Actions Speak Louder than Words: Environmental Law Enforcement and Audit Fees. *Review of Accounting Studies*, 30(1), 519–574. <https://doi.org/10.1007/s11142-024-09823-x>
- Zelevnikow, J. (2023). The Benefits and Dangers of Using Machine Learning to Support Making Legal Predictions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 13(4). <https://doi.org/10.1002/widm.1505>