



## Transparansi Algoritma Prediktif dalam Proses Peradilan Pidana: Analisis Kesenjangan antara Akurasi Teknis dan Hak Asasi Manusia

Dhani Prasetyo<sup>1\*</sup>, Elsa Sofiana Rahma<sup>2</sup>

<sup>1,2</sup> Universitas YARSI, Indonesia

Alamat: Jl. Letjend Suprpto, Cempaka Putih, Kota Administrasi Jakarta Pusat, Daerah Khusus Jakarta

Email : [Dhaniprasetyo@yarsi.ac.id](mailto:Dhaniprasetyo@yarsi.ac.id)<sup>1</sup>, [Elsasofianarahma@yarsi.ac.id](mailto:Elsasofianarahma@yarsi.ac.id)<sup>2</sup>

Korespondensi penulis: [Dhaniprasetyo@yarsi.ac.id](mailto:Dhaniprasetyo@yarsi.ac.id)

**Abstract:** *The application of predictive algorithms in criminal justice processes, such as for recidivism risk assessment or bail decisions, promises increased efficiency and objectivity. However, algorithmic transparency becomes a critical issue when confronted with human rights. This study analyzes the gap between the technical accuracy of predictive algorithms and the protection of human rights (HAM), particularly the right to a fair trial, the presumption of innocence, and the right to know and understand the basis of decisions affecting oneself. Using a normative-juridical approach and comparative study of practices in several jurisdictions, this research finds that a focus on statistical accuracy often neglects procedural transparency and accountability. Even technically accurate algorithms can violate human rights if their internal workings are "black boxes," inaccessible to defendants, or non-contestable. The findings reveal a structural gap between technical metrics (e.g., prediction accuracy rates) and compliance with human rights standards. Therefore, a regulatory framework is needed that mandates meaningful transparency, the right to explanation, and independent oversight of the use of predictive algorithms in criminal courts.*

**Keywords:** *predictive algorithms, transparency, criminal justice, human rights, technical accuracy.*

**Abstrak :** Penerapan algoritma prediktif dalam proses peradilan pidana, seperti untuk penilaian risiko residivis atau penentuan jaminan, menjanjikan peningkatan efisiensi dan objektivitas. Namun, transparansi algoritma menjadi isu krusial ketika berhadapan dengan hak asasi manusia (HAM). Penelitian ini menganalisis kesenjangan antara akurasi teknis algoritma prediktif dan perlindungan HAM, khususnya hak atas peradilan yang adil, presumsi tak bersalah, dan hak untuk mengetahui dan memahami dasar keputusan yang mempengaruhi diri seseorang. Melalui pendekatan yuridis-normatif dan studi komparatif atas praktik di beberapa yurisdiksi, studi ini menemukan bahwa fokus pada akurasi statistik sering mengabaikan prinsip transparansi prosedural dan akuntabilitas. Algoritma yang akurat secara teknis sekalipun dapat melanggar HAM apabila mekanisme kerja internalnya bersifat "kotak hitam" (black box), tidak dapat diakses oleh terdakwa, atau tidak dapat digugat. Hasil penelitian menunjukkan kesenjangan struktural antara metrik teknis (misalnya rasio prediksi benar) dan kepatuhan terhadap standar HAM. Oleh karena itu, diperlukan kerangka regulasi yang mewajibkan transparansi bermakna, hak interpretasi, dan pengawasan independen atas penggunaan algoritma prediktif di peradilan pidana.

**Kata kunci:** algoritma prediktif, transparansi, peradilan pidana, hak asasi manusia, akurasi teknis.

### 1. PENDAHULUAN

Perkembangan kecerdasan buatan (AI) dan pembelajaran mesin (machine learning) telah merambah ke berbagai sektor kehidupan, termasuk sistem peradilan pidana. Di berbagai negara, aparat penegak hukum dan hakim mulai menggunakan algoritma prediktif untuk membantu keputusan penting seperti penentuan risiko pelanggaran ulang (residivis), penetapan besaran jaminan, keputusan pembebasan bersyarat, hingga rekomendasi hukuman. Algoritma-algoritma ini, seperti COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) di Amerika Serikat, HART di Inggris, atau berbagai sistem serupa yang mulai dikembangkan

di negara lain, menjanjikan objektivitas, konsistensi, dan efisiensi yang lebih tinggi dibandingkan penilaian subjektif manusia yang rentan terhadap bias.

Namun, di balik janji akurasi teknis tersebut, muncul persoalan mendasar yang mengancam hak asasi manusia (HAM). Algoritma prediktif sering kali beroperasi sebagai "kotak hitam" (black box) karena kompleksitasnya, penggunaan data historis yang mungkin mengandung bias diskriminatif, serta perlindungan kekayaan intelektual dari pengembang perangkat lunak. Akibatnya, transparansi algoritma menjadi sangat terbatas. Terdakwa, penasihat hukum, bahkan hakim sekalipun, sering tidak dapat mengetahui secara pasti faktor-faktor apa saja yang digunakan algoritma untuk menghasilkan skor risiko tertentu, bagaimana bobot masing-masing faktor, atau apakah algoritma tersebut telah memproses data secara adil.

Kondisi ini menciptakan kesenjangan (gap) yang serius antara dua nilai yang seharusnya seimbang: akurasi teknis versus perlindungan HAM. Secara teknis, sebuah algoritma mungkin dianggap "akurat" jika tingkat prediksi kebenarannya tinggi (misalnya, benar memprediksi siapa yang akan melakukan tindak pidana ulang). Namun, akurasi teknis tidak menjamin kepatuhan terhadap prinsip-prinsip HAM, seperti hak atas peradilan yang adil (fair trial), hak presumption of innocence ( praduga tak bersalah), hak untuk didengar, hak untuk mengetahui dan memahami dasar keputusan yang mempengaruhi dirinya (right to explanation), serta hak untuk mengajukan keberatan (right to contest).

Sebagai contoh, kasus terkenal *State v. Loomis* (2017) di Wisconsin, AS, memperlihatkan bahwa algoritma COMPAS yang digunakan hakim untuk menghukum terdakwa tidak dapat diperiksa secara terbuka karena diklaim sebagai rahasia dagang. Pengadilan memutuskan bahwa penggunaan algoritma tersebut tidak melanggar due process, tetapi keputusan ini menuai kritik luas karena mengorbankan transparansi dan hak terdakwa untuk mengetahui dasar putusan yang menjatuhnya.

Di Indonesia, meskipun penggunaan algoritma prediktif dalam peradilan pidana belum sistematis, tren digitalisasi peradilan melalui e-court, e-berkas, dan pengembangan "smart court" menunjukkan bahwa adopsi AI dalam proses peradilan adalah keniscayaan. Tanpa kajian kritis dan kerangka regulasi yang memadai, Indonesia berisiko mengulang kesalahan yang sama: mengutamakan akurasi teknis di atas perlindungan HAM.

Oleh karena itu, penelitian ini penting untuk dilakukan guna menganalisis secara mendalam kesenjangan antara akurasi teknis algoritma prediktif dan hak asasi manusia dalam proses peradilan pidana, serta merumuskan rekomendasi kebijakan yang menyeimbangkan efisiensi teknologi dengan perlindungan hak-hak fundamental warga negara.

## 2. KAJIAN TEORITIS

### 2.1 Teori Transparansi Algoritma (*Algorithmic Transparency Theory*)

Transparansi algoritma merujuk pada sejauh mana cara kerja, logika, data masukan, dan proses pengambilan keputusan suatu algoritma dapat diakses, dipahami, dan diperiksa oleh pihak-pihak yang terkena dampaknya, termasuk pengguna, subjek data, dan pengawas publik. Menurut Pasquale (2015), algoritma yang digunakan dalam pengambilan keputusan publik cenderung menjadi "black box" karena kerahasiaan komersial, kompleksitas teknis, dan kurangnya kewajiban hukum untuk membukanya. Dalam konteks peradilan pidana, transparansi algoritma mencakup tiga dimensi: (a) transparansi kode sumber, (b) transparansi data pelatihan, dan (c) transparansi proses pengambilan keputusan individual (explainability). Tanpa transparansi yang memadai, mustahil untuk menilai apakah suatu algoritma telah bekerja secara adil dan tidak diskriminatif.

### 2.2 Teori Akurasi Teknis dalam Sistem Prediktif (*Technical Accuracy Theory*)

Akurasi teknis dalam algoritma prediktif biasanya diukur menggunakan metrik seperti true positive rate, false positive rate, precision, recall, dan \*F1-score\*. Dalam konteks peradilan pidana, algoritma prediktif sering dievaluasi berdasarkan seberapa tepat algoritma memprediksi perilaku masa depan terdakwa, misalnya risiko residivis. Namun, teori ini mengakui adanya accuracy-fairness trade-off, yakni peningkatan akurasi untuk satu kelompok dapat menurunkan akurasi atau keadilan untuk kelompok lain. Akurasi teknis juga tidak secara otomatis menjamin validitas konstruk, yaitu apakah variabel yang diukur (misalnya "risiko") benar-benar mencerminkan realitas yuridis yang relevan. Selain itu, akurasi historis (berdasarkan data masa lalu) tidak menjamin akurasi prospektif di masa depan, apalagi jika kondisi sosial berubah.

### 2.3 Teori Hak atas Peradilan yang Adil (*Right to a Fair Trial*)

Dalam instrumen HAM internasional seperti Pasal 10 DUHAM, Pasal 14 ICCPR, dan hukum nasional seperti UUD 1945 Pasal 28D ayat (1), hak atas peradilan yang adil mencakup hak untuk diadili oleh pengadilan yang independen dan imparial, hak atas proses yang terbuka, hak untuk mengetahui dakwaan, hak untuk didengar, hak untuk mengajukan bukti, dan hak untuk mengajukan banding. Dalam konteks algoritma prediktif, penerapan teori ini menuntut bahwa setiap keputusan yang dipengaruhi oleh algoritma harus tetap dapat dipertanggungjawabkan kepada hakim dan terdakwa. Algoritma tidak boleh menggantikan diskresi hakim secara absolut, dan terdakwa berhak untuk mempertanyakan serta menantang hasil keluaran algoritma sebagai bagian dari proses kontradiksi.

## 2.4 Prinsip Akuntabilitas dan Due Process of Law

Prinsip due process of law (proses hukum yang adil) mengharuskan bahwa setiap keputusan yang membatasi hak individu harus didasarkan pada prosedur yang jelas, diketahui sebelumnya, dan memberikan kesempatan untuk mengajukan keberatan. Dalam kaitannya dengan algoritma, teori procedural regularity menekankan bahwa proses pembuatan keputusan algoritmik harus memenuhi standar prosedural yang sama dengan keputusan manusia, termasuk kewajiban memberikan alasan (reason-giving requirement). Dworkin (1971) dalam pandangannya tentang substantive due process juga menekankan bahwa prosedur tidak boleh secara substansial tidak adil, meskipun secara formal tampak netral. Akuntabilitas algoritmik membutuhkan adanya mekanisme gugatan (contestability), audit independen, serta identifikasi entitas yang bertanggung jawab ketika algoritma menghasilkan keputusan yang keliru.

## 2.5 Teori Kesenjangan Akurasi-HAM (*Accuracy-Rights Gap Theory*)

Teori ini dikembangkan sebagai kerangka untuk mengidentifikasi situasi di mana peningkatan performa teknis suatu sistem tidak sejalan dengan atau bahkan merugikan perlindungan HAM. Kesenjangan dapat terjadi dalam berbagai bentuk: (a) epistemic gap ketidakmampuan pemangku kepentingan untuk memahami bagaimana akurasi dicapai; (b) procedural gap tidak tersedianya mekanisme untuk menguji atau menggugat hasil algoritma yang akurat secara teknis namun tidak adil secara individual; (c) remedial gap tidak adanya sanksi atau koreksi ketika algoritma yang akurat secara teknis melanggar HAM. Kesenjangan ini diperparah jika sistem hukum belum mengakomodasi pembuktian diskriminasi algoritmik karena beban pembuktian yang berat dan sifat "kotak hitam" algoritma.

## 3. METODE PENELITIAN

### 3.1 Jenis Penelitian

Penelitian ini menggunakan jenis penelitian hukum normatif (*doctrinal legal research*), yaitu penelitian yang mengkaji hukum sebagai sistem norma, asas-asas, aturan perundang-undangan, dan doktrin hukum. Pendekatan ini dipilih karena fokus penelitian adalah menganalisis kesenjangan antara standar teknis akurasi algoritma prediktif dan norma-norma hak asasi manusia yang mengikat dalam proses peradilan pidana.

### 3.2 Pendekatan Penelitian

**Tabel 1 Pendekatan Penelitian**

| Pendekatan                                       | Keterangan                                                                                                                                                                                                                                                         |
|--------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Pendekatan Perundang-undangan (Statute Approach) | Mengkaji berbagai peraturan perundang-undangan nasional dan internasional yang relevan, seperti KUHAP, UU HAM, UU ITE, serta instrumen HAM internasional (ICCPR, UDHR).                                                                                            |
| Pendekatan Kasus (Case Approach)                 | Menganalisis putusan pengadilan, baik dari yurisdiksi asing (misalnya <i>State v. Loomis</i> 2017) maupun yurisdiksi nasional yang relevan, untuk melihat bagaimana pengadilan memperlakukan bukti atau rekomendasi berbasis algoritma.                            |
| Pendekatan Komparatif (Comparative Approach)     | Membandingkan regulasi dan praktik di beberapa negara (misalnya AS, Inggris, Jerman, Estonia) yang telah mengadopsi algoritma prediktif dalam peradilan pidana untuk mengidentifikasi praktik terbaik (best practice) dalam menyeimbangkan akurasi teknis dan HAM. |
| Pendekatan Konseptual (Conceptual Approach)      | Mengkaji konsep-konsep hukum seperti transparansi, akuntabilitas, due process of law, praduga tak bersalah, dan keadilan algoritmik.                                                                                                                               |

### 3.3 Jenis dan Sumber Bahan Hukum

#### a. Bahan Hukum Primer

Bahan hukum primer yang digunakan meliputi:

Hukum Nasional Indonesia:

- Undang-Undang Dasar Negara Republik Indonesia Tahun 1945 (khususnya Pasal 28D dan 28I)
- Kitab Undang-Undang Hukum Acara Pidana (KUHAP)
- Undang-Undang No. 39 Tahun 1999 tentang Hak Asasi Manusia
- Undang-Undang No. 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik beserta perubahannya
- Peraturan Mahkamah Agung tentang e-court dan digitalisasi peradilan

Hukum Internasional:

- Universal Declaration of Human Rights (UDHR) 1948
- International Covenant on Civil and Political Rights (ICCPR) 1966
- UNESCO Recommendation on the Ethics of Artificial Intelligence (2021)

- d. Council of Europe's Framework Convention on Artificial Intelligence and Human Rights (draft)

Regulasi asing (untuk perbandingan):

- a. EU's General Data Protection Regulation (GDPR) – khususnya hak atas penjelasan (right to explanation)
- b. EU's AI Act (2024) – khusus bab tentang high-risk AI systems di sektor peradilan pidana
- c. Guidelines dari OECD, FAT/CP (Council of Europe)

b. Bahan Hukum Sekunder

Bahan hukum sekunder meliputi buku-buku teks, artikel jurnal hukum, laporan penelitian, dan dokumen kebijakan yang membahas:

- a. Algoritma prediktif dalam peradilan pidana
- b. Transparansi dan akuntabilitas AI
- c. Hak asasi manusia dalam era digital
- d. Keadilan algoritmik (algorithmic fairness)

c. Bahan Hukum Tersier

Bahan hukum tersier seperti kamus hukum, ensiklopedia hukum, dan indeks literatur untuk memaknai istilah-istilah teknis dan yuridis.

### 3.4 Teknik Pengumpulan Bahan Hukum

Teknik yang digunakan adalah studi kepustakaan (library research). Bahan hukum dikumpulkan melalui:

- a. Penelusuran database hukum online (misalnya: Google Scholar, JSTOR, HeinOnline, SSRN)
- b. Penelusuran peraturan perundang-undangan melalui website resmi (BPK, Kemenkumham, UN Treaty Collection)
- c. Penelusuran putusan pengadilan asing melalui database seperti CourtListener, Harvard Law Review, atau Westlaw
- d. Studi dokumen dan laporan kebijakan dari lembaga internasional (UN, UNESCO, OECD, Council of Europe)

### 3.5 Teknik Analisis Bahan Hukum

Analisis data dalam penelitian ini menggunakan metode interpretasi hukum dan analisis kualitatif normatif, dengan langkah-langkah sebagai berikut:

- a. Reduksi data: Memilah dan mengkategorikan bahan hukum berdasarkan topik (transparansi, akurasi, HAM, kesenjangan).

Interpretasi:

- Interpretasi gramatikal (makna harfiah teks hukum)
- Interpretasi sistematis (hubungan antar pasal dan antar peraturan)
- Interpretasi teleologis (tujuan perlindungan HAM di balik suatu norma)
- Analisis kesenjangan (gap analysis): Membandingkan apa yang diwajibkan oleh norma HAM dengan apa yang dihasilkan atau diizinkan oleh standar teknis akurasi algoritma.
- Sintesis menyusun temuan menjadi suatu kerangka utuh tentang kesenjangan akurasi-HAM.
- Penarikan kesimpulan menggunakan logika deduktif (dari norma umum ke kasus konkret algoritma) dan induktif (dari studi kasus ke generalisasi normatif).

## 4. HASIL DAN PEMBAHASAN

### 4.1 Hasil Penelitian

#### 4.1.1 Akurasi Teknis Algoritma Prediktif: Antara Janji dan Realitas

Algoritma prediktif dalam peradilan pidana, seperti COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*) di Amerika Serikat, dirancang untuk menilai risiko residivis, membantu keputusan jaminan, pembebasan bersyarat, dan rekomendasi hukuman. Secara teknis, algoritma ini menggunakan puluhan hingga ratusan variabel (COMPAS menggunakan 137 variabel) untuk menghasilkan skor risiko.

**Tabel 2 Temuan mengenai akurasi teknis**

| Metrik                             | Temuan                                                                                                                                                                                                                                                                                          |
|------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Validitas Prediktif                | Algoritma menunjukkan tingkat akurasi yang bervariasi. Studi menunjukkan COMPAS memiliki akurasi sekitar 65% dalam memprediksi residivis, tidak jauh lebih baik dari prediksi manusia awam .                                                                                                    |
| Bias Algoritmik                    | Studi ProPublica (2016) menemukan COMPAS dua kali lebih mungkin salah mengklasifikasikan terdakwa kulit hitam sebagai berisiko tinggi ( <i>false positive</i> ) dibandingkan terdakwa kulit putih, yang lebih sering salah diklasifikasikan sebagai berisiko rendah ( <i>false negative</i> ) . |
| Fairness dalam Alokasi Sumber Daya | Penelitian kuantitatif di Bogotá dan Chicago menunjukkan bahwa model prediktif yang adil dalam prediksi belum tentu adil dalam alokasi patroli. Perbedaan rata-rata manfaat antar kelompok mencapai 12-45%                                                                                      |

#### 4.1.2 Transparansi Algoritma: The Black Box Problem

Salah satu temuan paling signifikan adalah kurangnya transparansi algoritma prediktif yang digunakan di pengadilan pidana.

Kasus *State v. Loomis* (2016) menjadi tonggak penting. Pengadilan Wisconsin mengizinkan penggunaan COMPAS meskipun algoritma tersebut bersifat proprietary dan tidak dapat diperiksa oleh terdakwa karena diklaim sebagai rahasia dagang. Keputusan ini menuai kritik karena mengorbankan hak terdakwa untuk mengetahui dasar keputusan yang mempengaruhi dirinya.

Temuan tentang transparansi:

- a. Black box problem: Kompleksitas algoritma (137 variabel pada COMPAS) membuat mustahil untuk melacak bagaimana bobot diberikan pada setiap faktor.
- b. Kerahasiaan dagang: Pengembang perangkat lunak sering melindungi kode sumber sebagai rahasia dagang, menghalangi pemeriksaan independen.
- c. Kesenjangan prosedural: Terdakwa tidak memiliki mekanisme yang bermakna untuk menggugat atau mempertanyakan hasil algoritma

### 4.2 Pembahasan

#### 4.2.1 Ketegangan antara Akurasi dan Keadilan Prosedural

Temuan di atas menunjukkan adanya ketegangan fundamental antara dua nilai yang seharusnya seimbang dalam sistem pengadilan pidana: akurasi teknis dan keadilan prosedural.

Secara teknis, algoritma dapat dianggap "akurat" jika tingkat prediksi kebenarannya tinggi. Namun, akurasi teknis tidak menjamin kepatuhan terhadap prinsip *due process of law*. Sebagaimana dijelaskan dalam literatur, keadilan prosedural bukanlah sinonim dari akurasi keadilan prosedural menuntut agar proses hukum memperlakukan subjek dengan martabat, memberikan kesempatan untuk didengar, dan memastikan adanya akuntabilitas.

#### 4.2.2 Transparansi vs. Rahasia Dagang: Konflik yang Tak Terelakkan

Salah satu temuan paling kritis adalah konflik antara kebutuhan transparansi dalam proses pengadilan dan perlindungan kekayaan intelektual (rahasia dagang) oleh pengembang algoritma.

Dalam sistem pengadilan yang adversarial, proses kontradiksi (*adversarial process*) adalah mekanisme utama untuk menguji kebenaran. Namun, ketika alat algoritmik dilindungi sebagai rahasia dagang, mekanisme ini terhambat:

- a. Terdakwa tidak dapat mengetahui bagaimana sistem bekerja
- b. Tidak dapat mengidentifikasi mode kegagalan yang diketahui
- c. Tidak dapat menilai apakah alat tersebut telah divalidasi secara independen.

## 5. KESIMPULAN DAN SARAN

### 5.1 Kesimpulan

#### 5.1.1 Terdapat Kesenjangan Fundamental antara Akurasi Teknis dan Perlindungan HAM

Penelitian ini menemukan bahwa kesenjangan antara akurasi teknis algoritma prediktif dan perlindungan hak asasi manusia bersifat sistemik dan struktural, bukan sekadar masalah teknis yang dapat diatasi dengan penyempurnaan algoritma semata. Kesenjangan tersebut terwujud dalam tiga bentuk utama:

- a. Kesenjangan epistemik (Epistemic Gap) : Pemangku kepentingan (terdakwa, penasihat hukum, bahkan hakim) tidak mampu memahami bagaimana suatu algoritma mencapai tingkat akurasi tertentu karena sifat "kotak hitam" (black box) dan kompleksitas teknis yang tidak dapat ditembus.
- b. Kesenjangan prosedural (Procedural Gap) : Tidak tersedianya mekanisme hukum yang bermakna untuk menguji, mempertanyakan, atau menggugat hasil keluaran algoritma dalam proses peradilan yang bersifat adversial.
- c. Kesenjangan remedial (Remedial Gap) : Tidak adanya sanksi hukum atau mekanisme koreksi yang efektif ketika algoritma yang secara teknis akurat ternyata melanggar hak asasi manusia, seperti yang terjadi dalam kasus *State v. Loomis*.

#### 5.1.2 Akurasi Teknis Tidak Menjamin Kepatuhan terhadap *Due Process of Law*

Penelitian ini membuktikan bahwa akurasi teknis yang tinggi yang diukur melalui metrik seperti true positive rate, false positive rate, AUC, dan kalibrasi tidak secara otomatis menjamin kepatuhan terhadap prinsip *due process of law*, praduga tak bersalah (*presumption of innocence*), hak atas peradilan yang adil (*fair trial*), dan hak untuk mengetahui serta menggugat dasar keputusan yang mempengaruhi diri seseorang (*right to explanation and contestation*). Sebaliknya, fokus berlebihan pada optimalisasi akurasi statistik justru cenderung mengabaikan dimensi keadilan prosedural yang merupakan inti dari sistem peradilan pidana yang beradab.

#### 5.1.3 Konflik antara Transparansi dan Rahasia Dagang Tidak Terelakkan

Terdapat konflik yang melekat dan tidak terelakkan antara tuntutan transparansi dalam proses peradilan pidana dan perlindungan kekayaan intelektual (rahasia dagang) yang diklaim oleh pengembang algoritma. Dalam kasus *State v. Loomis*, pengadilan mengizinkan penggunaan algoritma COMPAS meskipun kode sumbernya tidak dapat diperiksa oleh terdakwa karena dilindungi sebagai rahasia dagang. Keputusan ini mengorbankan hak fundamental terdakwa untuk mengetahui dasar putusan yang menjeratnya. Solusi yang ditawarkan bukanlah membuka seluruh kode sumber secara publik, melainkan mewajibkan

empat kategori informasi minimal yang diperlukan untuk keadilan prosedural: validasi independen, tingkat kesalahan terdisagregasi, ruang lingkup operasional, dan akuntabilitas manusia.

#### **5.1.4 Algoritma Prediktif Berisiko Melanggengkan dan Memperkuat Bias Historis**

Temuan dari kasus COMPAS dan berbagai yurisdiksi lain (Spanyol, Belanda, dll.) menunjukkan bahwa algoritma prediktif yang dilatih dengan data historis akan mereproduksi dan bahkan memperkuat bias yang sudah ada dalam data tersebut. Data historis penangkapan dan hukuman sering mencerminkan praktik diskriminatif masa lalu (misalnya, over-policing terhadap kelompok minoritas). Algoritma tidak memiliki kemampuan untuk membedakan antara "risiko sebenarnya" dan "risiko yang tercatat karena diskriminasi sistemik". Hal ini bertentangan secara fundamental dengan asas praduga tak bersalah yang mengharuskan setiap individu diperlakukan berdasarkan perilaku dan bukti individual, bukan berdasarkan karakteristik kelompok atau data historis yang bias.

#### **5.1.5 Indonesia Belum Siap Mengadopsi Algoritma Prediktif tanpa Regulasi yang Memadai**

Penelitian ini menemukan bahwa Indonesia masih berada dalam tahap awal diskusi mengenai penggunaan kecerdasan buatan di peradilan. Beberapa hakim telah mulai menggunakan AI dalam putusan, namun belum ada regulasi spesifik yang mengatur transparansi, akuntabilitas, dan perlindungan HAM. Terdapat kesenjangan regulasi yang serius: KUHAP yang ada tidak mengakomodasi tantangan pembuktian algoritmik, tidak ada standar audit HAM untuk AI di peradilan, dan belum ada pedoman tentang hak terdakwa untuk menolak atau menggugat keputusan berbasis algoritma. Tanpa kerangka hukum yang memadai, Indonesia berisiko mengulang kesalahan yang sama seperti yang terjadi di Amerika Serikat dan yurisdiksi lain.

### **5.2 Saran**

#### **5.2.1 Untuk Pemerintah dan Pembuat Kebijakan**

- a. Pertama, pemerintah harus segera menyusun regulasi khusus tentang penggunaan kecerdasan buatan (AI) di peradilan pidana. Regulasi ini wajib mengatur kewajiban transparansi algoritma, hak terdakwa untuk mendapatkan penjelasan, serta mekanisme untuk menggugat hasil keputusan algoritma.
- b. Kedua, sebelum regulasi yang memadai terbentuk, pemerintah perlu memberlakukan moratorium atau penghentian sementara terhadap penggunaan algoritma prediktif yang bersifat "kotak hitam" dalam keputusan yang berdampak besar pada hak individu, seperti penahanan, penentuan jaminan, atau pemidanaan.

- c. Ketiga, jika tetap ingin melakukan uji coba, uji coba tersebut harus dilakukan dalam skala terbatas, diawasi secara independen oleh lembaga HAM nasional dan akademisi, serta hasil evaluasinya dipublikasikan secara terbuka.
- d. Keempat, Mahkamah Agung harus memastikan bahwa setiap kebijakan digitalisasi peradilan (seperti e-court dan smart court) menyertakan standar perlindungan hak asasi manusia sebagai syarat utama, bukan sekadar pelengkap.

### **5.2.2 Untuk Hakim dan Aparat Penegak Hukum**

- a. Pertama, hakim tidak boleh sepenuhnya menyerahkan keputusan kepada algoritma. Algoritma hanya boleh berfungsi sebagai alat bantu, bukan pengganti kewenangan hakim. Keputusan akhir harus tetap berada di tangan manusia yang bertanggung jawab.
- b. Kedua, jika algoritma digunakan, hakim harus memastikan bahwa terdakwa dan penasihat hukumnya mendapatkan akses informasi yang cukup tentang cara kerja algoritma, variabel yang digunakan, dan tingkat kesalahan yang terdokumentasi.
- c. Ketiga, hakim harus tegas menolak penggunaan hasil algoritma yang tidak dapat diperiksa, tidak memiliki validasi independen, atau tidak disertai mekanisme kontestasi yang bermakna.

### **5.2.3 Untuk Pengembang Teknologi dan Perusahaan**

- a. Pertama, untuk algoritma yang digunakan dalam keputusan publik yang berdampak pada hak asasi manusia, perlindungan rahasia dagang tidak dapat dijadikan alasan untuk menolak pemeriksaan independen. Pengembang harus bersedia membuka algoritma untuk diaudit publik.
- b. Kedua, pengembang wajib menyediakan dokumentasi yang jelas dan transparan, meliputi data pelatihan yang digunakan, variabel beserta bobotnya, tingkat kesalahan per kelompok demografis, hasil validasi independen, serta batasan penggunaan yang direkomendasikan.
- c. Ketiga, sistem AI untuk peradilan pidana harus dirancang sejak awal agar dapat diaudit, transparan, dan memiliki mekanisme kontestasi yang terintegrasi.

### **5.2.4 Untuk Akademisi dan Peneliti**

- a. Pertama, peneliti perlu mengembangkan metode audit yang dapat menguji apakah suatu algoritma prediktif sudah memenuhi standar hak asasi manusia, termasuk pengukuran bias, transparansi, dan akuntabilitas.
- b. Kedua, diperlukan penelitian empiris di Indonesia untuk memahami persepsi hakim, jaksa, dan advokat terhadap penggunaan AI di peradilan, serta kesiapan dan risiko yang mungkin timbul.

- c. Ketiga, para akademisi harus mengkaji ulang secara mendalam konsep *due process of law* di era digital, terutama bagaimana hak untuk didengar, hak untuk mengetahui, dan hak untuk menggugat harus diadaptasi dalam sistem peradilan yang menggunakan keputusan algoritmik.

#### **5.2.5 Untuk Masyarakat Sipil dan Lembaga HAM**

- a. Pertama, Komnas HAM dan organisasi masyarakat sipil harus secara aktif mendorong pemerintah untuk membentuk regulasi yang melindungi hak asasi manusia dalam penggunaan AI di peradilan pidana.
- b. Kedua, lembaga HAM perlu membangun kapasitas teknis untuk memantau dan mengevaluasi penggunaan algoritma prediktif di lembaga peradilan dan kepolisian.
- c. Ketiga, masyarakat sipil harus meningkatkan literasi publik tentang risiko penggunaan AI dalam peradilan pidana, termasuk hak-hak warga negara ketika berhadapan dengan sistem pengambilan keputusan algoritmik.

#### **5.2.6 Untuk Penelitian Selanjutnya**

- a. Pertama, peneliti selanjutnya disarankan melakukan studi kasus mendalam di negara-negara yang telah menggunakan algoritma prediktif untuk mempelajari dampak nyata terhadap hak-hak terdakwa.
- b. Kedua, diperlukan penelitian interdisipliner yang melibatkan ilmuwan komputer, ahli hukum, dan aktivis HAM untuk merancang sistem AI yang adil dan transparan.
- c. Ketiga, peneliti juga dapat mengkaji pengalaman negara-negara Asia Tenggara lainnya seperti Malaysia, Singapura, dan Thailand untuk mengidentifikasi praktik terbaik yang sesuai dengan konteks Indonesia.

#### **DAFTAR REFERENSI**

- Corbett-Davies, S., & Goel, S. (2018). The Measure and Mismeasure of Fairness: A Critical Review of Fair Machine Learning. arXiv preprint, arXiv:1808.00023.
- Ervitasari, D., & Mayasari, R. T. (2026). The Void of Artificial Intelligence Criminal Accountability Norms in Indonesian Criminal Law Reform. *Jurnal Studi Islam*, Universitas Muhammadiyah Bengkulu.
- Ferguson, A. G. (2017). *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. New York University Press.
- Hildebrandt, M. (2015). *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology*. Edward Elgar Publishing.

- Krastev, M. (2026). Artificial Intelligence in Criminal Justice: Predictive Tools, Evidentiary Challenges and Human Rights Implications. *Faculty of Law Journal, Goce Delcev University*, 31(1), 1-19.
- Levendowski, A. (2018). How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem. *Washington Law Review*, 93, 579.
- Lewandowska, K., Reinhardt, L., & Whitmore, D. (2023). The Jurisprudence of Big Data: Legal Limits of Predictive Analytics in Criminal Justice Risk Assessment Models. *Legal Studies in Digital Age*.
- Liptrap, A. (2021). The Missing Algorithm: Safeguarding Brady Against the Rise of Trade Secrecy in Policing. *Michigan Law Review*, 119(6).
- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group.
- Oswalt, K. (2025). Algorithmic Sentencing: Balancing Efficiency with Due Process Rights. *AI Law - International Review of Artificial Intelligence Law*.
- Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.
- Roth, A. (2016). Trial by Machine. *Georgetown Law Journal*, 104, 1245.
- State v. Loomis, 881 N.W.2d 749 (Wis. 2016), cert. denied, 137 S. Ct. 2290 (2017).
- Supratman Andi Agtas. (2026). Draft Presidential Rule to Regulate AI Use. *ANTARA News*.
- Vollmer, S., & Matzner, T. (2025). Algorithmic and Non-Algorithmic Fairness: Should We Revise our View of the Latter Given Our View of the Former? *Law and Philosophy*, 44, 155-179.
- Wexler, R. (2018). Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System. *Stanford Law Review*, 70, 1343.
- Zavrsnik, A. (2019). Algorithmic Justice: Algorithms and Big Data in Criminal Justice Settings. *European Journal of Criminology*, 18(5), 623-641.