BAB II

KAJIAN PUSTAKA

2.1 Penelitian Terdahulu

Penelitian dalam bidang deteksi anomali konsumsi energi telah mengalami evolusi signifikan, dari pendekatan reaktif klasik menuju pengembangan sistem berbasis teknologi. Namun, studi mendalam terhadap literatur yang ada mengungkap kesenjangan mendasar: sebagian besar penelitian terjebak dalam paradigma "*detection-centric*" yang mengoptimalkan akurasi algoritma tanpa mengatasi tantangan mendasar terkait verifikasi dan kepercayaan.

Korpus penelitian saat ini dapat diklasifikasikan ke dalam tiga kelompok utama: (1) solusi berbasis *machine learning* yang berfokus pada identifikasi pola, (2) implementasi *blockchain* untuk integritas data, dan (3) sistem IoT untuk pemantauan *real-time*. Meskipun setiap kluster menunjukkan perkembangan teknis yang pesat, ketiadaan integrasi holistik dan metodologi verifikasi yang terdesentralisasi menciptakan kesenjangan konseptual yang besar.

Penelitian sebelumnya dalam kluster ***Machine Learning* untuk Deteksi Anomali** menunjukkan fokus yang kuat pada pengoptimalan akurasi deteksi. Zhuang et al., 2024 mengadopsi *Graph Convolutional Networks* (GCN) dengan *residual learning* untuk meningkatkan akurasi deteksi, tetapi strategi ini mengabaikan bagaimana temuan deteksi dapat dikonfirmasi secara independen oleh banyak pemangku kepentingan. Cai et al., 2023 mengembangkan hibridisasi *Random Forest* dengan *Support Vector Data Description* (SVDD) yang menunjukkan peningkatan dalam deteksi anomali berbasis pembelajaran statistik, tetapi tidak menyediakan mekanisme untuk menerjemahkan hasil deteksi menjadi bukti yang dapat diverifikasi.

Wu & Wu, 2024 mengusulkan model CNN-BiLSTM hibrida dengan mekanisme *self-attention* yang menunjukkan akurasi prediksi lebih baik, tetapi intensitas komputasinya yang tinggi membuatnya tidak cocok untuk penerapan di lingkungan *edge*. Bahkan lebih signifikan lagi, strategi ini tetap berada dalam struktur pemrosesan terpusat yang tidak mengatasi asimetri kepercayaan dasar antara pemangku kepentingan energi.

Analisis terhadap penelitian terdahulu secara konsisten menyoroti sebuah *trade-off* kritis yang belum terpecahkan dalam konteks deteksi anomali energi di lingkungan IoT. Salah satu temuannya menunjukkan evolusi algoritma *gradient boosting* yang mengarah pada superioritas LightGBM untuk aplikasi IoT energy monitoring. Hipotesis penelitian ini: LightGBM akan menunjukkan kinerja superior dibandingkan algoritma *machine learning* lainnya untuk deteksi anomali konsumsi energi dalam lingkungan IoT yang *resource-constrained*. Penelitian oleh Ke et al., 2017 memperkenalkan paradigma baru dalam *gradient boosting* melalui *leaf-wise tree growth strategy*. Berbeda dengan pendekatan *level-wise* yang digunakan XGBoost T. Chen & Guestrin, 2016, LightGBM memilih *leaf* dengan *delta loss* terbesar untuk ekspansi yang menghasilkan *loss reduction*.

Di satu sisi, model *deep learning* seperti CNN-BiLSTM Wu & Wu, 2024 menunjukkan akurasi prediksi yang tinggi. Literatur mengakui bahwa tuntutan komputasi yang signifikan dan persyaratan memori yang besar dari model-model ini sering kali membuatnya tidak layak untuk diimplementasikan pada perangkat *edge* yang sumber dayanya terbatas, kesulitan yang didokumentasikan secara ekstensif dalam penelitian *edge intelligence* (J. Chen et al., 2023). Sebaliknya, sistem saat ini sering kali tidak menggabungkan deteksi waktu nyata dengan mekanisme verifikasi yang efektif, seperti yang terlihat pada metodologi seperti FWHT Civelek et al., 2024 atau *Random Forest* & KNN (Taruna et al., 2025), yang terus bergantung pada intervensi manual atau kurang koneksi IoT *real-time*. Kesenjangan ini menguraikan serangkaian kriteria ketat untuk model optimal dalam kerangka TDAVF:

1. Mampu mendeteksi pola anomali yang rumit, mirip dengan model canggih.
2. Memiliki latensi inferensi yang sangat rendah dan ukuran memori yang kecil untuk bekerja secara efektif pada perangkat *edge*.
3. Mampu mengelola volume data IoT yang besar dan cepat secara efektif.

Dalam spektrum algoritma *machine leanring*, keluarga model *Gradient Boosting Decision Tree* (GBDT) muncul sebagai kandidat utama yang berpotensi mampu memenuhi persyaratan yang saling bertentangan ini. Secara khusus, *Light Gradient Boosting Machine* (LightGBM), yang diusulkan oleh (Ke et al., 2017), dirancang secara arsitektural untuk mengatasi keterbatasan efisiensi yang terlihat pada implementasi GBDT lainnya. Berbagai studi perbandingan telah mengkonfirmasi bahwa manfaat desain LightGBM secara rutin menghasilkan waktu pelatihan dan inferensi yang lebih

cepat dibandingkan dengan GBDT terkemuka lainnya seperti XGBoost, terutama pada kumpulan data yang luas (Bentéjac et al., 2021). Manfaat ini muncul dari arsitektur khasnya dengan alasan sebagai berikut.

1. Algoritma Berbasis Histogram. Mengurangi penggunaan memori dan mempercepat proses pelatihan.
2. Pertumbuhan Tree Per Leaf. Memungkinkan konvergensi lebih cepat dan seringkali menghasilkan akurasi yang lebih baik.
3. Teknik Pengambilan Sampel yang Efisien. Menerapkan *Gradient-based One-Side Sampling* (GOSS) dan *Exclusive Feature Bundling* (EFB) secara signifikan mengurangi kompleksitas komputasi.

Lebih penting lagi, rekam jejak penggunaan LightGBM yang terbukti di bidang-bidang terkait mendukung posisinya sebagai pilihan utama. Penelitian oleh (Mbey et al., 2024) secara efektif menggunakan LightGBM untuk deteksi anomali dalam data jaringan pintar dengan akurasi tinggi dan latensi rendah. Demikian pula, dalam domain IoT Industri, keandalan LightGBM untuk tugas-tugas real-time juga telah ditunjukkan. (M. Wang et al., 2023) berhasil mengimplementasikan sistem diagnosis kesalahan berbasis IoT yang menggunakan LightGBM untuk mengidentifikasi anomali operasional secara akurat dan tepat waktu, membuktikan keseimbangan yang sangat baik antara kecepatan dan kinerja yang diperlukan untuk sistem pemantauan otonom.

Berdasarkan sintesis dari kesenjangan literatur, keunggulan arsitektural yang terbukti, dan keberhasilan penerapannya di domain yang relevan, penelitian ini merumuskan hipotesis utama: LightGBM adalah model klasifikasi yang paling optimal untuk kerangka kerja TDAVF, karena mampu memberikan kombinasi unggul antara akurasi deteksi, kecepatan inferensi waktu nyata, dan efisiensi komputasi untuk penerapan pada perangkat tepi. Hipotesis ini akan diuji secara empiris melalui analisis komparatif yang ketat sebagaimana diuraikan dalam Bab 3.

Kluster penelitian ***Blockchain* untuk Integritas Data** berfokus pada penggunaan teknologi *blockchain* untuk memastikan integritas data dalam transaksi energi. (Din et al., 2024) mempelajari *blockchain* untuk transaksi energi *peer-to-peer* yang aman, tetapi gagal mengatasi kesulitan latensi dan skalabilitas saat menangani aliran data IoT volume tinggi. (Z. Zhao et al., 2023) menciptakan deteksi anomali aman berbasis *blockchain* yang

menjamin penyimpanan data transparan, tetapi tidak menyediakan mekanisme bagi pengguna untuk mengelola akses data secara granular.

Penelitian dalam kluster **IoT untuk Pemantauan *Real-Time*** menekankan kemampuan sensor IoT untuk pemantauan konsumsi waktu nyata. (Z. Zhao et al., 2024) menerapkan deteksi pencurian yang menjaga privasi menggunakan enkripsi homomorfik, tetapi ketergantungan pada desain terpusat menimbulkan titik kegagalan tunggal dan kerentanan terhadap manipulasi. (Binyamin et al., 2024) menciptakan pemrosesan informasi *real-time* untuk mengenali pola pencurian listrik, namun volume data yang dihasilkan oleh perangkat IoT membentuk hambatan serius karena metode pemrosesan data konvensional tidak mampu mengikuti volume dan kompleksitas data secara *real-time*.

Analisis mendalam terhadap literatur mengungkapkan bahwa penelitian yang ada secara sistematis mengabaikan transformasi dari deteksi ke verifikasi. Mayoritas penelitian berfokus pada cara meningkatkan akurasi deteksi (Iftikhar, Khan, et al., 2024), tetapi mereka tidak menjawab pertanyaan mendasar: *bagaimana hasil deteksi dapat diubah menjadi bukti yang dipercaya secara universal dan dapat diverifikasi secara independen tanpa memerlukan otoritas pusat?*

Kesenjangan konseptual ini menciptakan asimetri kepercayaan yang persisten dalam ekosistem energi. Ketika anomali terdeteksi, bukti seringkali bersifat kepemilikan, dikendalikan oleh sistem utilitas, dan rentan terhadap skeptisisme dari sisi konsumen. Ketiadaan mekanisme bukti yang dapat diverifikasi secara kriptografis membuat para pemangku kepentingan energi terjebak dalam konflik, alih-alih memecahkan masalah secara kolaboratif.

(Civelek et al., 2024) menggunakan *Fast Walsh–Hadamard Transform* (FWHT) dengan *machine learning* untuk deteksi pencurian geospasial, tetapi respons yang tertunda terhadap peristiwa pencurian dan ketergantungan pada intervensi manusia menunjukkan tidak adanya protokol verifikasi otomatis. (Taruna et al., 2025) mengoptimalkan *Target Operation* (TO) menggunakan *Random Forest* & *K-Nearest Neighbors* (KNN) dengan akurasi 0,89, namun kurang integrasi IoT *real-time* dan mekanisme verifikasi bukti independen.

Berdasarkan studi literatur yang tela dilakukan, *Trust-Based Decentralized Anomaly Verification Framework* (TDAVF) memposisikan dirinya sebagai pergeseran paradigma

fundamental. Alih-alih mengoptimalkan akurasi deteksi dalam kerangka kerja terpusat, TDAVF mengatasi penyebab utama asimetri kepercayaan melalui pengenalan "*verified facts*" yang dapat diverifikasi secara kriptografis dan dapat diaudit secara independen oleh semua pemangku kepentingan energi. TDAVF mengatasi tiga kekurangan penting yang disoroti dalam literatur sebagai berikut.

1. **Kesenjangan Integrasi Arsitektur**. Kurangnya integrasi sistematis antara *edge intelligence*, konsensus *blockchain*, dan *adaptive machine learning* dalam *resource-constrained energy infrastructure*.
2. **Kesenjangan Protokol Kepercayaan**. Tidak adanya protokol kriptografi yang memungkinkan verifikasi pemangku kepentingan tanpa mengungkapkan data konsumsi sensitif.
3. **Kesenjangan Kerangka Skalabilitas**. Kurangnya *computational* dan *economic models* untuk menerapkan sistem verifikasi terdesentralisasi di berbagai lingkungan infrastruktur energi.

*Trust-Based Decentralized Anomaly Verification Framework (*TDAVF) membangun fondasi teknologinya berdasarkan konvergensi tiga kerangka teoretis yang sudah mapan dan telah divalidasi secara terpisah dalam domain komputasi dan sistem terdistribusi, dengan menggabungkannya menjadi satu arsitektur untuk menangani masalah spesifik dalam deteksi dan verifikasi penyalahgunaan energi.

Pertama, di bidang ***edge intelligence***, TDAVF mengembangkan ide *Gradient Boosting Decision Tree* (GBDT), yang terbukti unggul oleh (Ke et al., 2017) dalam kerangka LightGBM. Berbeda dengan pendekatan berbasis *deep learning* seperti CNN-LSTM (Nazmul Hasan et al., 2019) atau CNN-XGBoost, yang memiliki keterbatasan dengan *dataset* yang tidak seimbang (Nawaz et al., 2023), TDAVF mengoptimalkan LightGBM melalui pendekatan optimasi pada lingkungan *resource-constraints* (Boyd & Vandenberghe, 2014) untuk perangkat *edge* dengan memori terbatas. Dengan dasar ini, TDAVF mempertahankan akurasi yang tinggi dalam aplikasi *real-time* tanpa mengorbankan kinerja komputasi.

Kedua, untuk **generasi bukti yang *immutable***, TDAVF menyediakan arsitektur berdasarkan *Byzantine Fault Tolerance* (Lamport et al., 1982), yang merupakan landasan konsensus *blockchain* modern. Pendekatan ini mengintegrasikan bukti *Merkle Tree* (Merkle, 1987) dan tanda tangan digital (Rivest et al., 1978) pada perangkat *edge*,

memungkinkan setiap hasil deteksi anomali diterjemahkan menjadi fakta yang dapat diverifikasi dan disimpan secara permanen pada buku besar terdesentralisasi. Metode ini memastikan integritas data dan non-penyangkalan dalam ekosistem energi, melampaui implementasi *blockchain* sebelumnya yang masih bergulat dengan latensi IoT waktu nyata (Iqbal et al., 2023).

Ketiga, dalam hal **privasi dan kontrol data**, TDAVF membangun metodenya yang berbeda dengan skema terpusat yang rentan dengan memberdayakan pengguna melalui penerapan *Decentralized Identity Management Theory* dan konsep dari *Zero-Knowledge Proof Theory* (Goldwasser et al., 1985), yang memungkinkan pengguna untuk mempertahankan kontrol penuh atas data mereka. Kerangka kerja ini menggabungkan *Attribute-Based Access Control* (ABAC) *Theory (V. C. Hu et al., 2014)* dengan *smart contract* berbasis *blockchain*, menghasilkan sistem tokenisasi Web 3.0 di mana setiap otorisasi berbagi data diberikan secara kriptografis, tidak seperti sistem terpusat sebelumnya (Z. Zhao et al., 2024).

Kontribusi mendasar dari kerangka kerja ini bukan hanya pada akurasi deteksi, tetapi juga pada transformasi hasil deteksi menjadi bukti yang tidak dapat diubah dan dapat diverifikasi, serta pada mekanisme kontrol data yang lebih adil bagi pengguna dalam sistem energi digital.

2.2 Landasan Teori

2.2.1 Blockchain

Kemunculan teknologi *blockchain* menandai perubahan paradigma yang signifikan dalam evolusi sistem terdistribusi. Sebelum dirilisnya *white paper* Satoshi Nakamoto pada tahun 2008, kepercayaan dalam konteks digital didominasi oleh otoritas kelembagaan, server terpusat, atau pihak ketiga yang tepercaya atau *trusted third parties* (TTP). Perantara ini memvalidasi transaksi, menyelesaikan perselisihan, dan memastikan integritas catatan. Meskipun efisien secara fungsional, model-model ini rentan secara inheren: mereka menetapkan titik kegagalan tunggal, menimbulkan ketidakseimbangan kepercayaan antara penyedia layanan dan konsumen, dan membuat asal data bergantung pada kebaikan atau kemampuan otoritas pusat (Essaaidi et al., 2021).

Tindakan Nakamoto pada tahun 2008 secara fundamental mengubah epistemologi kepercayaan ini. *Blockchain* mengubah kepercayaan dari otoritas kelembagaan menjadi konsensus terdistribusi yang dapat diverifikasi secara kriptografis, memungkinkan agen

yang tidak saling percaya untuk berkolaborasi, bertransaksi, dan menjaga buku besar bersama tanpa wasit pusat. Terobosan ini mengatasi masalah pengeluaran ganda yang telah mengganggu iterasi sebelumnya dari mata uang digital dan, secara lebih umum, mengubah konsep pembentukan "*fact*" dalam sistem komputasi. Pada dasarnya, *blockchain* membangun konsep *trustless trust* yaitu kondisi di mana kebenaran dan integritas muncul bukan dari penegakan hierarkis tetapi dari interaksi primitif kriptografi dan koordinasi terdesentralisasi (Nakamoto, 2008).

Pada dasarnya, *blockchain* beroperasi sebagai buku besar terdesentralisasi yang disebarkan melalui jaringan *peer-to-peer* (P2P). Setiap *node* menyimpan salinan seluruh riwayat transaksi, menjamin bahwa tidak ada satu entitas pun yang dapat memonopoli atau memodifikasi data tanpa terdeteksi. *Blockchain* adalah serangkaian blok yang diatur secara berurutan, masing-masing berisi kumpulan transaksi yang diverifikasi. Transaksi sering diatur dalam pohon *Merkle*, memfasilitasi jalur verifikasi cepat untuk kumpulan data yang luas. Setiap blok memiliki referensi hash kriptografi yang merujuk pada pendahulunya, *timestamp*, dan *nonce* (Sasikumar et al., 2024). Ketergantungan siklis blok memastikan tidak dapat diubah: bahkan satu pembaruan tunggal pada data historis menyebar ke depan, membatalkan semua blok berikutnya (Dong et al., 2023).

Teknik konsensus menentukan bagaimana *node* menyetujui standar dari buku besar. *Proof of Work* (PoW), meskipun menuntut komputasi dan memberikan tekanan ekologis, menetapkan mekanisme yang kuat untuk melindungi jaringan dari serangan Sybil dengan mengikat konsensus pada upaya komputasi (Nakamoto, 2008). Peningkatan selanjutnya seperti *Proof of Stake* (PoS) memperkenalkan opsi yang lebih hemat energi, menurunkan jejak ekologis operasi *blockchain* lebih dari 99% sambil meningkatkan throughput secara bersamaan (H. Guo & Yu, 2022). Di luar konsensus, *public-key cryptography* mendasari validitas transaksi: peserta menggunakan kunci pribadi untuk menghasilkan tanda tangan digital, yang kemudian dapat diverifikasi menggunakan kunci publik yang sesuai, sehingga menetapkan kepemilikan dan mengotorisasi transfer nilai tanpa pernah mengungkap *private key* itu sendiri (Krichen et al., 2022).

Kegunaan *blockchain* sejak itu telah jauh melampaui *bitcoin*. Munculnya *smart contracts,* kode yang dieksekusi sendiri yang disebarkan di rantai mengubah *blockchain* menjadi infrastruktur yang dapat diprogram, mampu menggabungkan logika rumit langsung ke dalam buku besar (Din et al., 2024). Terobosan ini telah memungkinkan

aplikasi baru seperti perdagangan energi *peer-to-peer*, keuangan terdesentralisasi, penelusuran rantai pasokan yang transparan, dan bahkan sistem deteksi anomali yang mencatat perilaku mencurigakan secara permanen (Z. Zhao et al., 2023). Teknologi penskalaan komplementer, termasuk solusi Layer-2 seperti *sidechain* dan *state channel*, telah mengurangi beberapa batasan throughput dan latensi yang melekat pada *blockchain* lapisan dasar, sehingga kasus penggunaan *real-time* dan frekuensi tinggi menjadi semakin praktis (Neiheiser et al., 2023).

Namun, meskipun ada perkembangan konseptual dan teknis ini, evaluasi cermat terhadap *blockchain* dalam konteks *Internet of Things* (IoT) dan sistem energi mengungkap batasan substansial yang belum dibahas dalam literatur.

1. **Ketidakselarasan Waktu**. *Blockchain*, bahkan dengan algoritma konsensus yang ditingkatkan, beroperasi pada periode konfirmasi blok yang diukur dalam detik atau menit. Sebaliknya, perangkat IoT sering menghasilkan data dengan interval milidetik atau subdetik, terutama dalam infrastruktur vital seperti distribusi listrik. Ketidakcocokan waktu membuat *blockchain* tidak cocok sebagai substrat pengambilan keputusan *real-time* (Iqbal et al., 2023).
2. **Kendala Biaya dan Efisiensi**. Menyimpan setiap pembacaan sensor atau peristiwa konsumsi sebagai transaksi *on-chain* menciptakan biaya yang sangat tinggi, baik dalam upaya komputasi maupun biaya transaksi ("gas"). Dalam skala besar, hal ini menyebabkan beban ekonomi dan komputasi yang tidak berkelanjutan.
3. ***Garage In, Garbage Out* (GIGO)**. *Blockchain* memberikan ketidakmungkinan perubahan catatan setelah diserahkan, tetapi tidak dapat menjamin kebenaran input mentah. Jika sensor yang rusak atau cacat menghasilkan data yang salah, blockchain akan dengan setia menyimpan kesalahan tersebut secara permanen. Kekurangan ini sangat signifikan di sektor-sektor di mana integritas data sangat penting untuk deteksi anomali (Finck, 2019).
4. ***Residual Trust Asymmetry* pada *Permissioned Systems***. Banyak solusi *blockchain* untuk sistem industri atau energi menggunakan model berizin (*permissioned model*) untuk meningkatkan efisiensi. Namun, ini menambahkan otoritas manajemen pusat yang mengatur partisipasi, mengkompromikan premis desentralisasi dan mengembalikan asimetri dalam kepercayaan (Z. Zhao et al.,

2023).

5. **Ketidakcocokan Arsitektur dengan *Edge Processing***. Ekosistem IoT semakin mengandalkan komputasi *edge* untuk analitik yang sensitif terhadap latensi. Ketergantungan *blockchain* pada konsensus global di semua *node* secara fundamental tidak selaras dengan tuntutan pemrosesan lokal dan waktu nyata dari lingkungan *edge*, menghasilkan kesenjangan struktural antara auditabilitas dan responsivitas (Din et al., 2024).

Secara bersamaan, kendala-kendala ini menunjukkan kontradiksi yang belum terpecahkan pada penerapan *blockchain* untuk pemantauan energi berbasis IoT: *blockchain* menawarkan integritas dan auditabilitas pasca-peristiwa, tetapi gagal memberikan verifikasi pra-peristiwa atau responsivitas *real-time*. Solusi saat ini menekankan ketidakmampuan berubahnya *blockchain* dengan mengorbankan efisiensi, atau memprioritaskan analitik *edge* yang cepat dan terpusat dengan mengorbankan auditabilitas dan kepercayaan.

Limitasi ini menghadirkan kesenjangan penelitian dalam aplikasi yang memerlukan *real-time data provenance* dan verifikasi integritas, di mana persyaratan pemrosesan *edge* bertentangan dengan *delay* konsensus yang melekat pada sistem *blockchain*. Literatur sebelumnya tidak memiliki paradigma arsitektural yang dapat menjembatani hal ini.

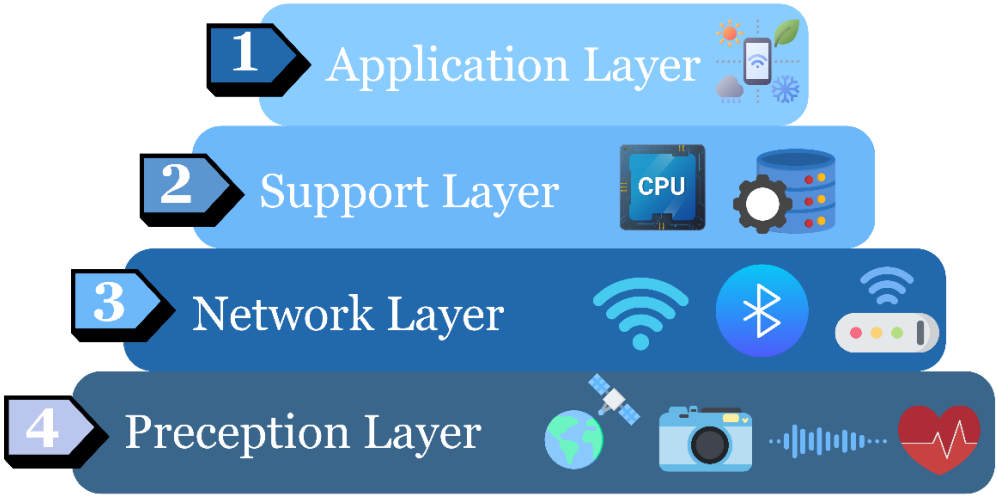
Lebih spesifik lagi, tidak ada *framework* yang mengatasi bagaimana mentransformasi hasil deteksi anomali *real-time* menjadi "verified facts" yang dapat diaudit secara independen oleh semua *stakeholder* tanpa memerlukan otoritas pusat, sambil mempertahankan efisiensi komputasi yang diperlukan untuk *deployment* pada perangkat *resource-constrained*. Gap ini menciptakan *persistent trust asymmetry* dalam ekosistem energi, di mana bukti anomali tetap bersifat proprietari dan rentan terhadap skeptisisme, bukannya menjadi fakta matematis yang dapat diverifikasi secara universal.

2.2.2 Internet of Things (IoT)

Internet of Things (IoT), sebuah konsep yang pertama kali diartikulasikan oleh Kevin Ashton pada tahun 1999, menggambarkan jaringan global objek fisik yang saling terhubung, atau "*Things*", yang disematkan dengan sensor, aktuator, dan kemampuan komunikasi untuk memfasilitasi pertukaran data dan pengambilan keputusan yang cerdas tanpa campur tangan manusia secara langsung (Ashton, 1999). *Internet of Things* (IoT) memperluas internet konvensional dengan memungkinkan benda-benda fisik untuk

mengumpulkan dan bertukar data sendiri, sehingga menghasilkan jaringan gadget berjejaring yang berinteraksi dan bekerja sama secara *real time* (Elgabri et al., 2023). Perubahan paradigma ini membuka pintu bagi berbagai aplikasi di berbagai sektor mulai dari rumah dan kota pintar hingga otomasi industri dan perawatan kesehatan.

Potensi IoT untuk mengintegrasikan dunia fisik dan digital menyediakan kerangka kerja untuk mengumpulkan, menganalisis, dan mengeksploitasi data untuk meningkatkan efisiensi, produktivitas, dan pengalaman pengguna. Jaringan sensor, protokol komunikasi, *cloud computing*, dan *edge computing* adalah komponen utama yang memungkinkan pembangunan ekosistem yang ada di mana-mana dan saling terhubung (Vishwakarma & Das, 2021). Pertimbangan IoT mencakup keamanan data, privasi, dan standar untuk memungkinkan interoperabilitas yang lancar, mendukung terciptanya solusi kreatif yang merevolusi cara kita berinteraksi dengan lingkungan sekitar. Berbeda dengan penggunaan internet biasa, IoT memperluas koneksi di luar komputer dan ponsel pintar untuk mencakup berbagai hal fisik, termasuk apa pun mulai dari peralatan rumah tangga dan gadget yang dapat dikenakan hingga peralatan industri dan peralatan kesehatan (Hasan et al., 2022).



Gambar 2.1 Arsitektur Empat Lapisan pada Sistem *Internet of Things* (IoT).

Arsitektur yang mendasari sistem IoT memiliki banyak lapisan seperti yang diilustrasikan pada gambar 2.1, dimulai dengan lapisan persepsi di mana sensor mengumpulkan data dari lingkungan fisik-seperti arus listrik dalam aplikasi *smart grid* dan aktuator menjalankan perintah untuk memodifikasi lingkungan tersebut. Data ini dikirimkan melintasi lapisan jaringan, memanfaatkan berbagai protokol komunikasi, ke

lapisan pemrosesan yang memanfaatkan platform cloud untuk penyimpanan dan analisis data berskala besar, atau *edge computation* untuk pemrosesan dengan latensi rendah dan terlokalisasi di dekat sumber data (L. Yang & Shami, 2021). Inti dari IoT adalah penciptaan ekosistem jaringan di mana gadget berinteraksi dan bekerja sama untuk meningkatkan berbagai bagian dari kehidupan kita. Interkoneksi ini memungkinkan terciptanya *smart contract*, kota, otomasi industri, sistem perawatan kesehatan, dan penggunaan lainnya. Jaringan sensor yang mengumpulkan data dari dunia fisik, protokol komunikasi yang memungkinkan koneksi tanpa batas, komputasi awan untuk pemrosesan dan penyimpanan data terpusat, dan *edge computation* untuk pemrosesan data lokal di tingkat perangkat adalah komponen penting dari IoT (Paolone et al., 2022). Salah satu keuntungan terpenting dari IoT adalah kemampuannya untuk menghubungkan dunia fisik dan digital (Paolone et al., 2022). Sistem IoT memberikan wawasan yang signifikan tentang dunia fisik dengan mengumpulkan dan menganalisis data. Data ini kemudian dapat digunakan untuk meningkatkan efisiensi, produktivitas, dan pengalaman pengguna. Koneksi berkelanjutan antara dunia digital dan fisik ini mencirikan potensi transformasi IoT di berbagai sektor.

Namun, meskipun memiliki potensi transformatif, paradigma IoT yang umum sering didefinisikan sebagai "*sense-and-send*" mengungkapkan batasan struktural ketika diterapkan pada area-area penting seperti deteksi anomali energi. Dalam instalasi normal, perangkat IoT mengumpulkan data dan mentransfernya ke server pusat atau platform *cloud*, tempat analitik yang membutuhkan banyak komputasi dilakukan. Meskipun berguna untuk banyak kasus penggunaan, paradigma ini memperkenalkan latensi yang signifikan, kebutuhan bandwidth, dan kerentanan terhadap titik kegagalan pusat (J. Li et al., 2022). Dalam keadaan yang memerlukan respons *real-time*, seperti deteksi lonjakan konsumsi listrik yang cepat, penundaan ini dapat membuat sistem tidak efisien atau bahkan mematikan. Selain itu, ketergantungan pada arsitektur terpusat menciptakan kemacetan dan membuat sistem rentan terhadap risiko pemadaman, serangan siber, atau manipulasi di tingkat penyedia layanan, sehingga mengembalikan asimetri kepercayaan yang dirancang untuk dihilangkan oleh IoT (R. Liu et al., 2022).

Pembatasan lain berasal dari keragaman dan sifat perangkat IoT yang terbatas sumber dayanya. Banyak sensor dan perangkat *embedded* berfungsi dengan daya pemrosesan, memori, dan cadangan energi yang rendah (Qiu et al., 2020). Akibatnya, mereka tidak

memiliki daya untuk menjalankan algoritma deteksi anomali tingkat lanjut secara lokal, sehingga bergantung pada server jarak jauh. Ketergantungan ini memperburuk masalah latensi dan menghambat skalabilitas solusi IoT dalam sistem energi skala besar di mana ribuan sensor terus-menerus menghasilkan data frekuensi tinggi (S. Chen et al., 2019). Kekhawatiran keamanan dan privasi juga tetap mendesak: pengumpulan dan transmisi data sensitif secara terus-menerus meningkatkan risiko pengawasan, kebocoran data, dan manipulasi yang tidak sah, terutama ketika protokol standarisasi dan interoperabilitas tidak konsisten diterapkan di berbagai penerapan yang heterogen.

Meningkatnya pentingnya *edge computing* bertujuan untuk meringankan beberapa tantangan ini dengan memindahkan komputasi lebih dekat ke sumber data. Berbeda dengan paradigma *cloud* terpusat, *edge computing* menyebarkan kecerdasan di antara perangkat dan gerbang lokal, memungkinkan pemrosesan latensi rendah dan pengambilan keputusan yang lebih cepat (Moura et al., 2024). Namun bahkan di sini, masalah yang belum terselesaikan terus berlanjut. Meskipun lebih kompeten daripada sensor mentah, tetap memiliki keterbatasan yang membatasi kapasitas mereka untuk menjalankan model *machine learning* yang kompleks dalam skala besar (Iqbal et al., 2023). Selanjutnya, meskipun *edge computing* meningkatkan responsivitas, ia tidak secara inheren menjamin keandalan atau auditabilitas keputusan yang diambil, menciptakan masalah akuntabilitas dalam aplikasi penting seperti manajemen energi.

Analisis ini mengungkapkan kesenjangan penelitian di persimpangan IoT dan kepercayaan. Literatur sebelumnya telah menetapkan kelayakan IoT dalam menangkap dan menganalisis peristiwa fisik, dan arsitektur *cloud-edge* telah meningkatkan efisiensi dan responsivitas. Namun, belum ada paradigma yang digunakan secara luas yang secara bersamaan menjamin (1) responsivitas *real-time*, (2) efisiensi komputasi untuk perangkat terbatas, dan (3) keyakinan yang terverifikasi terhadap validitas anomali yang diidentifikasi. Implementasi IoT saat ini seringkali memenuhi dua dari persyaratan ini dengan mengorbankan yang ketiga: sistem terpusat memberikan kepercayaan tetapi tidak responsif, sistem *edge* menawarkan responsivitas tetapi tidak dapat diverifikasi kepercayaannya, dan model hibrida terkadang mengalami masalah dengan skalabilitas komputasi. Belum ada arsitektur terintegrasi di mana perangkat IoT tidak hanya dapat merasakan dan menyampaikan data, tetapi juga memvalidasi dan meningkatkan anomali menjadi fakta yang dikonfirmasi sebelum memasukkannya ke dalam catatan yang tidak

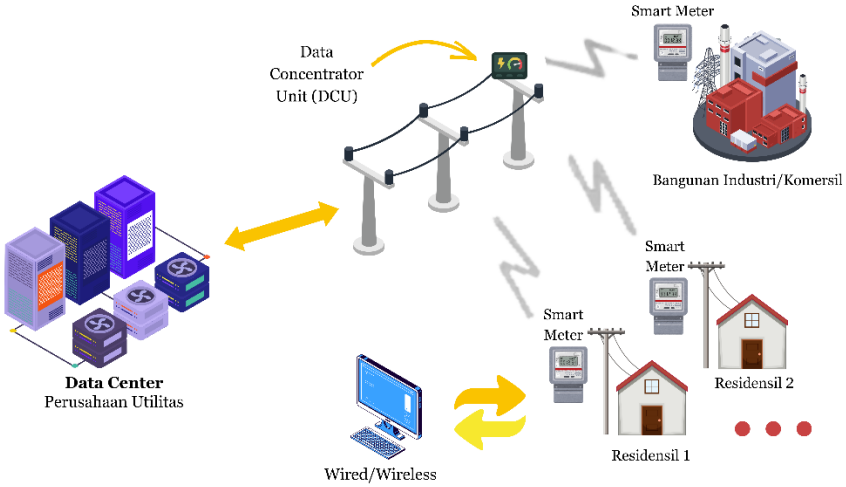
dapat diubah.

Di dalam celah yang belum terpecahkan inilah disertasi ini menempatkan kontribusinya. Dengan memikirkan kembali IoT tidak hanya sebagai jaringan sensor tetapi sebagai lapisan kecerdasan terdistribusi yang mampu verifikasi pra-rantai, penelitian ini bertujuan untuk mengatasi hambatan antara kecepatan dan kepercayaan. Perangkat IoT dan *node edge* berkolaborasi tidak hanya untuk mengumpulkan dan mengirimkan data, tetapi juga untuk memeriksa anomali secara lokal dengan cara terdesentralisasi, memastikan bahwa hanya fakta yang diverifikasi yang diteruskan untuk pencatatan permanen. Hal ini mengubah IoT dari sistem pengumpulan data pasif menjadi peserta aktif dalam pengembangan kepercayaan, menetapkan kerangka kerja untuk deteksi anomali yang tangguh, terukur, dan dapat diaudit dalam sistem konsumsi energi.

2.2.3 Smart Grid

Smart Grid melambangkan modernisasi besar-besaran dari jaringan tenaga listrik lama, beralih dari sistem terpusat dan searah menjadi jaringan cerdas, terdesentralisasi, dan dua arah melalui integrasi teknologi komunikasi dan informasi digital canggih. Meskipun konsep ini berkembang selama beberapa tahun, kerangka kerja dasar yang disediakan oleh lembaga-lembaga seperti *United States Department of Energy* (DOE) dan *National Institute of Standards and Technology* (NIST) mendefinisikannya sebagai jaringan pengiriman energi yang terdistribusi secara otomatis dan luas yang ditandai dengan aliran listrik dan informasi dua arah (Dorji et al., 2023). Arsitektur intinya dibangun di atas teknologi utama yang memungkinkan seperti *Advanced Metering Infrastructure* (AMI), yang mencakup smart meter untuk pelacakan konsumsi energi secara *real-time*, dan serangkaian sensor dan *Phasor Measurement Units* (PMU) yang digunakan di seluruh jaringan untuk memantau kondisinya dengan ketelitian tinggi (N. M. Kumar et al., 2020).

Struktur arsitektur canggih yang membedakannya dari jaringan listrik konvensional ini diilustrasikan pada Gambar 2.2. Lapisan teknis ini menyediakan pengumpulan data secara *real-time* dan memungkinkan kontrol dan otomatisasi yang canggih. Secara fungsional, sistem *smart grid* menyediakan penyeimbangan beban dinamis dan manajemen sisi permintaan, di mana utilitas dan pelanggan dapat terlibat untuk mengoptimalkan penggunaan energi berdasarkan harga waktu nyata dan keadaan sistem.



Gambar 2.2 Diagram Arsitektur *Smart Grid*.

Selain itu, sistem ini dimaksudkan untuk mengintegrasikan *Distributed Energy Resources* (DER) dengan mulus, seperti panel surya dan turbin angin, sehingga meningkatkan keberlanjutan dan fleksibilitas jaringan (Qays et al., 2023). Keuntungan penting dari paradigma ini adalah peningkatan ketahanannya; dengan memanfaatkan analitik waktu nyata dan peralihan otomatis, jaringan listrik dapat secara mandiri mendeteksi, mengisolasi, dan merespons kesalahan atau gangguan, sebuah kemampuan yang sering digambarkan sebagai "*self-healing*", yang secara signifikan meningkatkan keandalan dan mengurangi durasi pemadaman (Esenogho et al., 2022). Hamparan cerdas pada infrastruktur daya fisik ini mengubah jaringan menjadi sistem yang efisien, berkelanjutan, dan kuat yang mampu memenuhi kebutuhan energi modern.

Dalam implementasi AMI pada *smart grid*, ketersediaan data konsumsi beresolusi waktu memungkinkan analitik untuk memantau ketidakwajaran (*anomaly*) secara lebih dini. Namun, literatur menegaskan bahwa dalam ekosistem *Smart Grid*, anomali konsumsi listrik tidak bersifat tunggal, melainkan dapat muncul dalam beragam bentuk penyimpangan pola pemakaian terhadap kebiasaan historis, baik berupa lonjakan mendadak, pergeseran waktu penggunaan, maupun pola intermiten (Fenza et al., 2019). Keragaman bentuk anomali ini merupakan karakteristik umum pada data konsumsi energi yang dipengaruhi dinamika perilaku beban serta perubahan konteks pemakaian (Himeur et al., 2021).

Kompleksitas pola tersebut menggarisbawahi bahwa anomali tidak selalu identik dengan pencurian listrik atau *Non-Technical Loss* (NTL). Literatur NTL menunjukkan bahwa kehilangan non-teknis dapat berkaitan dengan *tampering*/*bypassing*, namun juga

dapat terkait *malfunctioning*/ketidakakuratan metering, kesalahan proses pencatatan/penagihan, atau faktor operasional lain (Glauner et al., 2017; Kim et al., 2019). Oleh karena itu, deteksi anomali dalam penelitian ini dipahami sebagai indikasi awal yang memicu kebutuhan verifikasi lanjutan, bukan sebagai kesimpulan final.

Lebih lanjut, tantangan analitik dalam domain ini diperketat oleh karakteristik data yang tidak seimbang (*Imbalanced*). Mengingat frekuensi kejadian anomali jauh lebih rendah dibandingkan pola konsumsi normal, berbagai studi merekomendasikan penerapan strategi penyeimbangan kelas (*class balancing*), seperti teknik *oversampling* atau *undersampling*, guna memitigasi bias model terhadap kelas mayoritas (Iftikhar, Khan, et al., 2024; Iftikhar, Mancha Gonzales, et al., 2024).

2.2.4 Primitif Kriptografi EVM

Pada implementasi *blockchain* berbasis *Ethereum Virtual Machine* (EVM), keamanan dan keterverifikasian data tidak hanya bergantung pada desain arsitektur jaringan, tetapi juga pada primitif kriptografi yang menjadi fondasi validasi transaksi dan integritas data. Dalam konteks penelitian ini, platform Polygon yang bersifat *EVM-compatible* digunakan, di mana primitif kriptografi utamanya mencakup fungsi hash Keccak-256 dan tanda tangan digital ECDSA.

a) **Fungsi *Hash* Keccak-256**

Keccak-256 merupakan fungsi *hash* kriptografis standar dalam ekosistem EVM. Pada level *smart contract*, fungsi keccak256 digunakan untuk menghasilkan *output* 256-bit yang unik dari input data tertentu (Solidity Documentation, 2023). Dalam protokol Ethereum, fungsi ini mendasari struktur data *State Trie*, di mana *stateRoot* pada *header* blok merupakan representasi *hash* dari seluruh status jaringan (Wood, 2014). Karakteristik ini memastikan integritas data: perubahan minimal pada input akan menghasilkan nilai hash yang berbeda secara signifikan, sehingga memudahkan pendeteksian rekayasa data.

b) **Digital Signature (ECDSA secp256k1)**

Tanda tangan digital berperan sebagai mekanisme otorisasi tanpa mengungkap kunci privat. Spesifikasi jaringan menggunakan *Elliptic Curve Digital Signature Algorithm* (ECDSA) pada kurva secp256k1 (Wood, 2014). Penggunaan kurva ini memungkinkan jaringan memverifikasi bahwa transaksi benar-benar berasal dari pemilik identitas yang sah, sekaligus menjamin prinsip *non-repudiation* di mana

pengirim tidak dapat menyangkal transaksi yang telah ditandatangani secara kriptografis (Mina Foundation, 2025).

c) **Immutability melalui *Hash Chaining***

Sifat *immutability* (ketidakberubahan) muncul dari keterikatan kronologis antarblok melalui ringkasan kriptografis. Setiap blok memuat *hash* dari blok sebelumnya, sehingga modifikasi pada data historis akan memicu ketidakkonsistenan pada seluruh rantai berikutnya. Hal ini menciptakan basis bukti yang permanen dan sulit dimanipulasi dalam jaringan terdistribusi.

Dalam kerangka penelitian ini, primitif tersebut diaplikasikan pada pengiriman bukti (*evidence*) dari perangkat IoT. Data anomali yang diidentifikasi oleh *edge gateway* (Raspberry Pi) diformat menjadi transaksi yang ditandatangani melalui ECDSA-secp256k1 sebelum disiarkan ke jaringan Polygon. Proses ini menjamin otentikasi pengirim dan memastikan bahwa bukti anomali tersimpan dengan integritas yang dapat diverifikasi secara independen oleh berbagai pemangku kepentingan (*multi-stakeholder*).

2.2.5 Auditabilitas & Sengketa Data

Anomali pada konsumsi listrik tidak selalu merepresentasikan satu penyebab tunggal. Dalam konteks sistem distribusi, anomali dapat mengindikasikan *non-technical losses* (NTL) seperti pencurian listrik, tetapi dapat pula disebabkan oleh faktor *non-malicious* seperti kerusakan/*defect metering*, gangguan perangkat, perubahan perilaku beban, maupun fenomena operasional lain pada jaringan. Literatur deteksi NTL menunjukkan bahwa tantangan utama bukan hanya mendeteksi pola yang menyimpang, tetapi juga mengelola ambiguitas penyebab anomali dan meningkatkan reliabilitas interpretasinya untuk kebutuhan operasional.

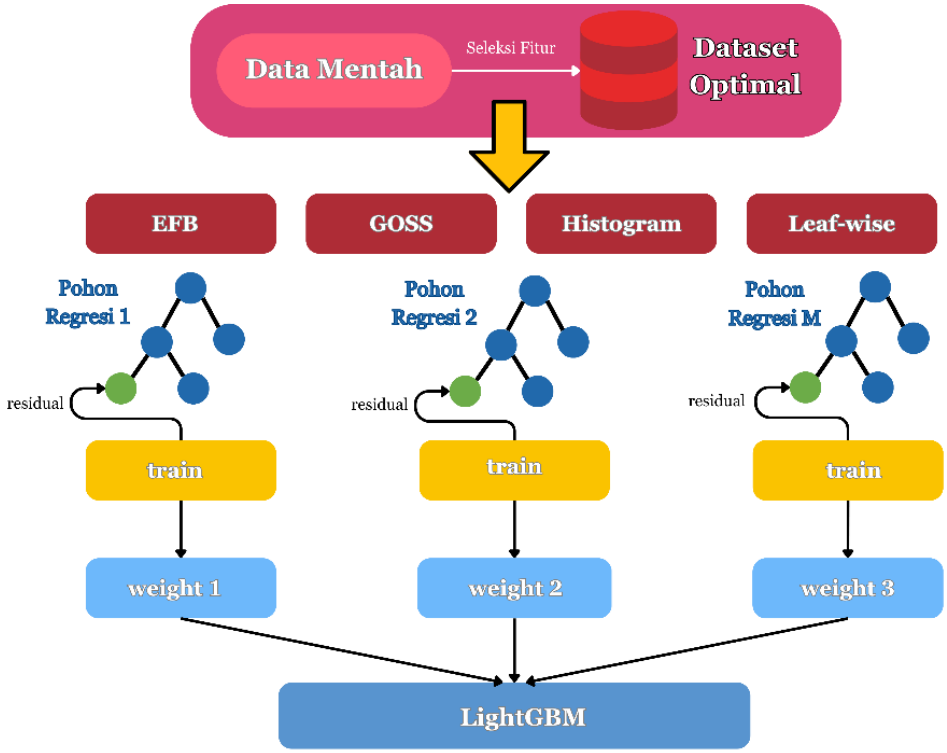
Sejumlah studi menegaskan keterkaitan anomali dengan dua kemungkinan besar: pencurian dan *defect metering*. Misalnya, kerangka deteksi anomali untuk mengendalikan NTL melaporkan bahwa koefisien/indikator anomali dapat mengarah pada pencurian energi *atau* cacat metering, sehingga memerlukan kehati-hatian dalam interpretasi hasil model (Yip et al., 2018). Di sisi lain, tinjauan sistematik dan review kontemporer NTL juga menekankan bahwa NTL mencakup variasi modus serangan/penyimpangan dan karakteristiknya, serta menunjukkan bahwa deteksi NTL kerap berhadapan dengan keterbatasan data dan problem implementasi di lapangan (Chuwa & Wang, 2021). Selain

keragaman pola, deteksi NTL sangat dipengaruhi oleh karakter data yang sering kali ***imbalanced*** (kasus pencurian lebih jarang dibanding konsumsi normal), sehingga berdampak pada bias model dan metrik evaluasi (Kim et al., 2019).

Sejumlah riset berbasis data konsumsi menyoroti penggunaan metode *machine learning* modern (termasuk *ensemble*) untuk memprediksi probabilitas NTL dan memperhatikan isu data konsumsi harian dalam skala besar. Studi lain (misal pada *Scientific Reports*) menunjukkan eksplorasi kerangka supervised learning untuk theft/NTL *detection* dan menekankan pentingnya metodologi yang relevan terhadap konteks implementasi (Abro et al., 2025). Dengan demikian, literatur menguatkan bahwa "anomali konsumsi" bersifat multi-penyebab dan memerlukan pendekatan yang tidak hanya mengejar akurasi deteksi, tetapi juga memperhatikan aspek interpretabilitas, verifikasi, dan bukti yang dapat digunakan lintas pihak (Coma-Puig & Carmona, 2022).

2.2.6 LightGBM

LightGBM adalah kerangka kerja peningkatan gradien yang menggunakan pohon keputusan sebagai pembelajar yang lemah dan dimaksudkan untuk menangani dataset berskala besar dan berdimensi tinggi secara efisien (Patel et al., 2020). LightGBM bekerja dengan menambahkan pohon baru secara berulang ke dalam model, dengan setiap pohon bertujuan untuk memperbaiki kesalahan yang dibuat oleh pohon sebelumnya (X. LightGBM menggunakan metode yang dikenal sebagai Gradient-based One-Side Sampling (GOSS) untuk mempersingkat waktu pelatihan sekaligus meningkatkan akurasi model (D. ni Wang et al., 2022). Karakteristik penting lainnya dari LightGBM adalah pendekatan pertumbuhan daun (*leaf-wise*), yang memungkinkan algoritma untuk memperluas kedalaman pohon terlebih dahulu, memilih simpul daun dengan penurunan terbesar dalam fungsi kerugian (J. Guo et al., 2023). Dibandingkan dengan kerangka kerja peningkatan gradien alternatif, metode ini dapat menghasilkan periode pelatihan yang lebih cepat dan akurasi yang lebih besar (Daoud, 2019).



Gambar 2.3 Diagram Algoritma LightGBM.

LightGBM menggunakan metode berbasis histogram untuk mengelompokkan informasi kontinu ke dalam nilai diskrit untuk lebih meningkatkan efisiensi model (Yan et al., 2021). Hal ini dapat meminimalkan penggunaan memori model dan waktu pelatihan, terutama untuk dataset berdimensi tinggi (Omotehinwa et al., 2023).

Penelitian terdahulu telah membuktikan akurasi yang luar biasa dalam mengidentifikasi anomali energi menggunakan model deep learning canggih (Wu & Wu, 2024). Meskipun ada kemajuan ini, pembatasan utama menghalangi penerapannya dalam konteks tepi dunia nyata yang terbatas sumber dayanya. Pertama, kompleksitas komputasi mereka sangat tinggi: Model CNN-BiLSTM dan GCN membutuhkan memori, daya CPU/GPU, dan waktu pemrosesan yang signifikan, sehingga tidak cocok untuk sistem seperti Raspberry Pi dengan RAM 1–4 GB. Latensi sering melebihi standar sub-detik untuk deteksi waktu nyata. Kedua, kebutuhan energi mereka tidak kompatibel dengan sensor IoT bertenaga baterai, karena operasi matriks berkelanjutan dengan cepat menghabiskan cadangan daya, mengganggu keberlanjutan skala besar. Ketiga, mereka berkinerja buruk pada kumpulan data yang tidak seimbang yang khas untuk deteksi anomali, sering kali lebih memilih konsumsi mayoritas (normal) sambil mengabaikan

peristiwa pencurian atau kegagalan yang tidak biasa tetapi signifikan (X. Guo et al., 2023). Terakhir, sifatnya yang "*black-box*" kurang memiliki kemampuan untuk dijelaskan seperti yang diperlukan dalam skenario di mana hasil anomali memengaruhi keputusan keuangan atau hukum, sehingga membatasi verifikasi yang transparan dan dapat diaudit.

Hal ini menunjukkan adanya kesenjangan penelitian yang mencolok. Literatur saat ini menuntut kompromi: model mendalam dengan akurasi tinggi tidak layak untuk penerapan di *edge*, atau model ringan mengorbankan presisi. Kesulitan semakin dalam ketika perangkat *edge* harus menyediakan bukti yang diverifikasi secara kriptografis, yang memerlukan deteksi anomali dan konstruksi bukti di bawah batasan sumber daya yang ketat. Yang dibutuhkan adalah kerangka algoritma yang khusus untuk verifikasi berbasis *edge* dan *real-time* mampu memproses aliran IoT berdimensi tinggi, mencapai presisi yang kuat pada kumpulan data yang tidak seimbang, menyediakan proses pengambilan keputusan yang transparan, dan berintegrasi dengan sistem bukti kriptografi. Hal ini menuntut perubahan paradigma menuju *machine leanrig* yang dioptimalkan untuk verifikasi, menyeimbangkan akurasi deteksi, efisiensi komputasi, interpretabilitas, dan integrasi *blockchain*. Pendekatan *ensemble* berbasis pohon, seperti *Lightweight Gradient Boosting* (LightGBM), muncul sebagai kandidat yang menarik: mereka dapat memproses data sensor dengan cepat, mempertahankan akurasi tinggi, tetap dapat ditafsirkan untuk audit trail, dan menyediakan verifikasi kriptografi tanpa bergantung pada sumber daya terpusat.

Tujuan utama dari algoritma LightGBM adalah untuk meminimalkan fungsi kerugian, yang merupakan penjumlahan dari kerugian individu untuk setiap contoh dalam dataset (X. Zhao et al., 2024). Tujuannya adalah untuk menemukan parameter model optimal θ yang meminimalkan fungsi objektif $L(\theta)$ seperti yang didefinisikan dalam Persamaan 2.1.

$$L(\emptyset) = \sum_{i=1}^{n} l\left(y_i, f(x_i, \emptyset)\right) + \Omega(\emptyset) \tag{2.1}$$

dengan:

$L(\theta)$ = Fungsi objektif (*objective function*),

$\sum_{i=1}^{n}$ = Penjumlahan atas seluruh instance data dalam himpunan pelatihan,

n = Kardinalitas dari himpunan data pelatihan,

$l(y_i, f(x_i; \theta))$ = Fungsi kerugian (*loss function*),

y_i = Nilai target observasi (*observed ground truth*) untuk contoh ke-i,

$f(x_i; \theta)$ = Fungsi prediksi (*prediction function*),

$\Omega(\theta)$ = Suku regularisasi sebagai penalisasi yang besarnya proporsional dengan kompleksitas model.

θ = Himpunan parameter model yang menjadi subjek optimisasi selama proses pelatihan.

Persamaan 2.1 merepresentasikan fungsi objektif yang menjadi target optimisasi dalam proses pelatihan model LightGBM. Fungsi ini tersusun atas dua komponen utama yang merefleksikan prinsip fundamental dalam pemodelan statistik, yaitu *bias-variance tradeoff* (T. Chen & Guestrin, 2016).

Suku Kerugian (*Loss Term*). Komponen pertama $\sum_{i=1}^{n} l(y_i, f(x_i, \emptyset)) + \Omega(\emptyset)$, dikenal sebagai risiko empiris (*empirical risk*). Suku ini mengkuantifikasi total deviasi atau ketidaksesuaian antara keluaran prediksi model, $f(x_i, \emptyset)$, dengan nilai target observasi, y_i, pada keseluruhan data pelatihan. Minimisasi suku ini bertujuan untuk meningkatkan akurasi dan ketepatan model terhadap data latih.

Suku Regularisasi (*Regularization Term*). Komponen kedua $\Omega(\theta)$, berfungsi sebagai penalisasi (*penalty*) terhadap kompleksitas model. Suku ini ditambahkan untuk mencegah fenomena *overfitting*, di mana model memiliki performa sangat tinggi pada data pelatihan namun gagal melakukan generalisasi dengan baik pada data yang belum pernah dilihat sebelumnya. Dengan membatasi kompleksitas, suku ini mendorong model untuk mempelajari pola yang lebih umum dan *robust*.

Tujuan memaksimalkan fungsi objektif ini adalah untuk menemukan kumpulan parameter terbaik, θ, yang mencapai keseimbangan yang tepat antara akurasi prediksi (dengan meminimalkan l*ost component*) dan *generalization capacity* (dengan mengelola suku regularisasi) (Islam et al., 2020).

Suku regularisasi $\Omega(\theta)$ pada LightGBM umumnya diformulasikan sebagai kombinasi dari regularisasi L1 dan L2, seperti yang tertera pada Persamaan 2.2.

$$\Omega(\emptyset) = \alpha \sum_{j=1}^{T} |w_j| + \frac{1}{2} \lambda \sum_{j=1}^{T} w_j^2 \quad (2.2)$$

dengan:

$\Omega(\theta)$ = Suku regularisasi total,
α = Koefisien regularisasi L1,
λ = Koefisien regularisasi L2,
T = Jumlah total terminal *nodes* (daun) dalam sebuah pohon keputusan,
w_j = Skor keluaran (bobot) yang diasosiasikan dengan leaf node ke-j,
$\sum_{j=1}^{T}$ = Agregasi (penjumlahan) atas seluruh *leaf nodes* dalam satu pohon,
w_j^2 = Kuadrat dari Norma L2 dari bobot w_j.

Persamaan 2.2 mendefinisikan mekanisme penalisasi kompleksitas model melalui bobot (*weights*) dari *leaf nodes* pada struktur *decision tree* (T. Chen & Guestrin, 2016). Setiap suku memiliki peran spesifik:

Regularisasi L1 (*Norm* L1). Suku $\alpha \sum |w_j|$ memberikan penalisasi berdasarkan magnitudo absolut dari bobot *leaf*. Efek utamanya adalah mendorong bobot dari *leaf* yang kurang kontributif menuju nol. Properti ini dikenal sebagai *sparsity* dan secara implisit berfungsi sebagai metode seleksi fitur, di mana model secara otomatis mengabaikan cabang-cabang pohon yang dianggap tidak signifikan.

Regularisasi L2 (*Norm* L2 Kuadrat). Suku $\frac{1}{2} \lambda \sum w_j^2$ memberikan penalisasi berdasarkan magnitudo kuadrat dari bobot. Teknik ini mencegah nilai bobot menjadi terlalu besar dan mendorong distribusi bobot yang lebih merata. Hal ini meningkatkan stabilitas model dan mengurangi sensitivitasnya terhadap fluktuasi kecil pada data pelatihan.

Kombinasi kedua teknik regularisasi ini (*Elastic Net regularization*) menyediakan mekanisme untuk mengontrol kompleksitas model, sehingga meningkatkan ketahanan (*robustness*) dan kemampuan generalisasinya. LightGBM menggunakan proses iteratif berbasis gradien untuk mengurangi fungsi objektif yang dijelaskan di atas. Arsitektur yang mengilustrasikan alur kerja dan komponen-komponen utama LightGBM ini

disajikan pada Gambar 2.3.

Model dimulai dengan prediksi dasar yang sederhana. Persamaan untuk inisialisasi model disajikan pada Persamaan 2.3 (Friedman, 2001).

$$F_0(x) = mean(y) \tag{2.3}$$

dengan:

$F_0(x)$ = Model Awal / Prediksi Dasar,
$mean(y)$ = Nilai Rata-rata Variabel Target.

Persamaan 2.3 menunjukkan langkah pertama dalam algoritma LightGBM, yaitu inisialisasi. Pada tahap ini, model awal ($F_0(x)$) tidak dilatih, melainkan langsung ditetapkan sebagai nilai rata-rata ($mean(y)$) dari seluruh variabel target pada data pelatihan. Nilai ini menjadi fondasi awal yang kemudian akan diperbaiki secara bertahap pada iterasi berikutnya.

Pada setiap iterasi, algoritma menghitung *error* residual dari model sebelumnya seperti yang ditunjukkan pada Persamaan 2.4.

$$r_{im} = y_i - F_{m-1}(x_i) \tag{2.4}$$

dengan:

r_{im} = Nilai Residual / Error,
y_i = Nilai Target Sebenarnya,
$F_{m-1}(x_i)$ = Prediksi Model Sebelumnya.

Setelah inisialisasi, model mulai belajar dari kesalahannya. Persamaan 2.4 menghitung nilai residual (r_{im}), yang merupakan selisih antara nilai target sebenarnya (y_i) dan prediksi yang dihasilkan oleh model pada iterasi sebelumnya ($F_{m-1}(x_i)$). Secara esensial, residual ini adalah representasi "*error*" dari model saat ini, dan nilai inilah yang menjadi target baru untuk dilatih oleh pohon keputusan berikutnya.

Setelah pohon baru dilatih, model *ensemble* diperbarui sesuai aturan pada Persamaan 2.5.

$$F_m(x) = F_{m-1}(x) + vh_m(x) \quad (2.5)$$

dengan:

$F_m(x)$ = Model Ensemble Terbaru,
$F_{m-1}(x)$ = Model Ensemble Sebelumnya,
v = Laju Pembelajaran (*Learning Rate*),
$h_m(x)$ = Pohon Keputusan Baru.

Persamaan 2.5 adalah aturan pembaruan model. Model ensemble terbaru ($F_m(x)$) dibentuk dengan mengambil model dari iterasi sebelumnya ($F_{m-1}(x)$) dan menambahkan pohon keputusan baru ($h_m(x)$) (Friedman, 2001). Kontribusi dari pohon baru ini diskalakan oleh laju pembelajaran (v), sebuah parameter kecil yang memastikan proses perbaikan model berjalan secara perlahan dan stabil untuk menghindari *overfitting*.

Setelah semua iterasi selesai, model final adalah penjumlahan dari model awal dan semua pohon yang telah dilatih pada Persamaan 2.6 .

$$\hat{y} = F_M(x) = \sum_{m=0}^{M} vh_m(x) \quad (2.6)$$

dengan:

$\hat{y}$ = Hasil Prediksi Final,
$F_M(x)$ = Fungsi Model Final,
$\sum_{m=0}^{M}$ = Penjumlahan dari model awal dan semua pohon keputusan yang telah dibangun,
$h_m(x)$ = Pohon Keputusan Baru.

Setelah sejumlah M iterasi selesai, Persamaan 2.6 digunakan untuk menghasilkan prediksi final ($\hat{y}$). Fungsi model final ($F_M(x)$) adalah hasil penjumlahan ($\sum$) dari model awal dan semua pohon keputusan ($h_m(x)$) yang telah dibangun. Setiap pohon memberikan kontribusi yang telah disesuaikan dengan laju pembelajaran, menghasilkan sebuah model akhir yang merupakan gabungan dari banyak 'pakar' lemah menjadi satu 'pakar' yang kuat.

2.2.7 Klasifikasi

Tugas klasifikasi pada *machine learning* didasarkan pada tujuan akhir membangun algoritma yang dapat secara otomatis mempelajari pola dan hubungan dalam data untuk mengklasifikasikan atau melabeli kejadian secara efektif ke dalam kelompok atau kategori yang telah ditentukan (Maxwell et al., 2018). Pelatihan model pada set data berlabel, di mana algoritma mempelajari pemetaan antara karakteristik input dan kelas output yang sesuai, adalah inti dari klasifikasi. Biasanya, teknik ini melibatkan pembelajaran yang diawasi, di mana model belajar dari contoh sebelumnya untuk membuat prediksi pada data baru yang sebelumnya tidak terlihat (Erkan, 2021). Ada banyak algoritma klasifikasi yang berbeda, mulai dari pendekatan konvensional seperti *decision tree*, *support vector machine*, dan regresi logistik hingga teknik yang lebih modern seperti jaringan syaraf dan metode ansambel seperti *random forest* (Alshammari, 2024).

Tugas klasifikasi didasarkan pada penyesuaian parameter model untuk mengurangi perbedaan antara label kelas yang diantisipasi dan label kelas yang sebenarnya. Ide generalisasi merupakan inti dari teori klasifikasi, di mana model yang dilatih menampilkan kinerja yang kuat pada data baru yang tidak diketahui, yang menunjukkan kapasitasnya untuk memprediksi pola di luar set pelatihan. Metrik evaluasi seperti akurasi, presisi, *recall*, dan skor F1 memberikan pengukuran kuantitatif untuk menganalisis keampuhan algoritma klasifikasi. Dasar-dasar teoritis kategorisasi *machine learning* berfungsi sebagai landasan untuk menangani berbagai masalah di dunia nyata, mulai dari diagnosis medis dan identifikasi gambar hingga penyaringan spam dan deteksi penipuan. Berikut merupakan persamaan metrik evaluasi akurasi, presisi, recall, dan skor F1 (Corral et al., 2025).

Evaluasi kuantitatif terhadap kinerja model klasifikasi secara fundamental didasarkan pada Confusion Matrix. Matriks ini menyajikan rekapitulasi hasil prediksi yang disandingkan dengan kelas aktualnya, dan terdekomposisi menjadi empat komponen utama:

a. **True Positive (TP):** Jumlah instans yang secara aktual berada di kelas positif dan berhasil diprediksi sebagai positif oleh model.
b. **True Negative (TN):** Jumlah instans yang secara aktual berada di kelas negatif dan berhasil diprediksi sebagai negatif.

c. **False Positive (FP):** Jumlah instans aktual negatif yang secara keliru diprediksi sebagai positif (Error Tipe I).
d. **False Negative (FN):** Jumlah instans aktual positif yang secara keliru diprediksi sebagai negatif (Error Tipe II).

Dari komponen-komponen ini, berbagai metrik evaluasi dapat diturunkan untuk menganalisis performa model dari berbagai perspektif.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.7}$$

dengan:

TP = *True Positive* (Benar Positif),
TN = *True Negative* (Benar Negatif),
FP = *False Positive* (Positif Palsu),
FN = *False Negative* (Negatif Palsu).

Akurasi merepresentasikan kapabilitas model secara keseluruhan dalam melakukan klasifikasi yang benar seperti yang disajikan pada Persamaan 2.7. Metrik ini mengkuantifikasi proporsi total instans, baik positif maupun negatif, yang diklasifikasikan dengan benar terhadap jumlah keseluruhan instans. Meskipun akurasi menyediakan gambaran umum yang intuitif mengenai performa model, utilitasnya dapat menurun secara signifikan pada skenario dengan distribusi kelas yang tidak seimbang (*imbalanced class distribution*), di mana metrik ini mungkin memberikan estimasi kinerja yang terlalu optimistik.

$$Precision = \frac{TP}{TP + FP} \tag{2.8}$$

dengan:

TP = *True Positive* (Benar Positif),
FP = *False Positive* (Positif Palsu).

Persamaan 2.8 menunjukkan Presisi, atau *positive predictive value* yang mengukur fraksi instans yang relevan (positif) di antara total instans yang telah diklasifikasikan sebagai positif. Metrik ini secara spesifik mengevaluasi reliabilitas dari prediksi positif yang dihasilkan oleh model. Tingkat presisi yang tinggi mengindikasikan rendahnya insiden *false positive* (FP). Metrik ini menjadi krusial dalam domain di mana konsekuensi

dari *false positive* signifikan, sehingga menuntut keyakinan tinggi pada setiap prediksi positif.

$$recall = \frac{TP}{TP + FN} \tag{2.9}$$

dengan:

TP = *True Positive* (Benar Positif),

FN = *False Negative* (Negatif Palsu).

Recall, yang juga dikenal sebagai sensitivitas atau *true positive rate* (TPR), mengukur kemampuan model untuk mengidentifikasi keseluruhan instans positif yang ada dalam dataset seperti pada Persamaan 2.9. Metrik ini mengevaluasi tingkat kelengkapan (*completeness*) dari prediksi positif yang dilakukan model. Tingkat recall yang tinggi mengindikasikan rendahnya insiden *false negative* (FN), yang menjadi prioritas utama dalam domain di mana kegagalan mendeteksi kasus positif memiliki konsekuensi yang berat.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{2.10}$$

dengan:

$Precision$ = Nilai Presisi dari Persamaan (2.8),

$Recall$ = Nilai *Recall* dari Persamaan (2.9).

Persamaan 2.10 menunjukkan F1-Score diformulasikan sebagai rata-rata harmonik (*harmonic mean*) dari presisi dan *recall*, yang menyediakan metrik tunggal untuk mengevaluasi keseimbangan antara kedua metrik tersebut. Dengan memperhitungkan baik *false positive* (melalui presisi) maupun *false negative* (melalui *recall*), F1-Score memberikan evaluasi performa yang lebih holistik, terutama pada kasus distribusi kelas yang tidak seimbang di mana akurasi saja tidak cukup representatif.

Berbeda dengan metrik berbasis *threshold* tunggal, AUC menyediakan evaluasi kinerja model secara agregat di seluruh spektrum *threshold* klasifikasi. Kurva ROC dibangun dengan memetakan *True Positive Rate* (TPR) terhadap *False Positive Rate* (FPR) seperti yang ditunjukkan pada Persamaan 2.11.

$$False\ Positive\ Rate\ (FPR) = \frac{FP}{FP + TN} \tag{2.11}$$

dengan:

TN = *True Negative* (Benar Negatif),

FP = *False Positive* (Positif Palsu).

Kurva ROC secara visual merepresentasikan *trade-off* diagnostik antara *true positive rate* (sumbu Y) dan *false positive rate* (sumbu X) pada berbagai tingkatan *threshold*. AUC (*Area Under the Curve*) mengkuantifikasi area total di bawah kurva ROC. Nilai AUC merepresentasikan probabilitas bahwa model akan memberikan peringkat (skor) yang lebih tinggi pada instans positif yang dipilih secara acak dibandingkan instans negatif yang dipilih secara acak. Oleh karena itu, AUC berfungsi sebagai metrik yang robust untuk mengukur daya diskriminatif model, independen dari *threshold* klasifikasi dan distribusi kelas.

2.2.8 Metode SMOTE

Synthetic Minority Over-sampling Technique (SMOTE) adalah pendekatan *machine learning* yang terkenal dan umum digunakan yang secara khusus dikembangkan untuk mengatasi masalah yang diberikan oleh set data yang tidak seimbang. SMOTE berasal dari sebuah teknik untuk mengatasi bias yang melekat pada dataset ketika satu kelas sangat mendominasi kelas lainnya, yang dikembangkan oleh Nitesh V. Chawla dan rekan-rekannya (Chawla et al., 2002). Asumsi utama SMOTE adalah pengembangan contoh kelas minoritas sintetis dengan melakukan interpolasi di antara contoh kelas minoritas yang ada. Strategi *oversampling* ini berusaha untuk memperbaiki ketidakseimbangan, sehingga meningkatkan kinerja prediksi model pembelajaran mesin, terutama dalam tugas klasifikasi (A. Li et al., 2023).

SMOTE tidak hanya mengurangi risiko bias model terhadap kelas mayoritas, tetapi juga mendiversifikasi set pelatihan, sehingga menghasilkan proses pembelajaran yang lebih kuat dan representatif (Guan et al., 2023). Keberhasilan SMOTE telah terbukti di berbagai aplikasi, memberikan kontribusi yang signifikan terhadap lanskap yang lebih besar dalam kategorisasi data yang tidak seimbang dan menekankan signifikansinya sebagai dasar teori untuk mengatasi kesulitan ketidakseimbangan kelas dalam *machine learning* (Mishra & Singh, 2021).

Pemanfaatan SMOTE menawarkan beberapa keuntungan, terutama dalam kemampuannya untuk mengatasi data yang tidak seimbang tanpa memerlukan upaya pengumpulan data tambahan. Dengan menghasilkan sampel sintetis, SMOTE secara

efektif memperluas dataset yang tersedia, memungkinkan pengklasifikasi untuk belajar dari representasi kelas yang lebih seimbang (Dai & Wang, 2019). SMOTE, sebuah teknik augmentasi data yang banyak digunakan, mengatasi masalah ini dengan menghasilkan sampel sintetis untuk kelas minoritas melalui interpolasi vektor fitur antara contoh-contoh yang berdekatan (B. Liu, 2023).

Proses augmentasi data dalam SMOTE tidak melibatkan duplikasi sederhana, melainkan generasi sampel sintetis baru melalui prosedur algoritmik yang terdefinisi. Untuk setiap instans pada kelas minoritas, algoritma ini melakukan serangkaian operasi untuk menghasilkan satu atau lebih instans sintetis. Prosedur ini dapat diuraikan sebagai berikut:

1. **Seleksi Instans Minoritas**: Sebuah instans, x_i, dipilih secara acak dari himpunan data kelas minoritas.
2. **Identifikasi Lingkungan *K-Nearest Neighbors***: Algoritma mengidentifikasi k-tetangga terdekat dari x_i yang juga merupakan anggota kelas minoritas. Parameter k adalah bilangan bulat yang dapat dikonfigurasi.
3. **Seleksi Acak Tetangga**: Satu tetangga, x_j, dipilih secara acak dari himpunan k-tetangga terdekat tersebut.
4. **Generasi Instans Sintetis**: Sebuah instans sintetis baru, dihasilkan melalui proses interpolasi linear antara x_i dan x_j berdasarkan Persamaan 2.12.

$$x_{new} = x_i + \lambda \times (x_j - x_i) \tag{2.12}$$

dengan:

x_{new} = Vektor Fitur Instans Sintetis,
x_i = Vektor Fitur Instans Minoritas Referensi,
x_j = Vektor Fitur Tetangga Minoritas Terpilih
λ = Skalar Acak pada Interval [0, 1]

Persamaan 2.12 merepresentasikan operasi interpolasi linear yang menjadi dasar generasi data sintetis pada SMOTE. Vektor fitur untuk instans sintetis baru (x_{new}) dihasilkan dengan mengambil vektor fitur instans minoritas referensi (x_i) dan menambahkan vektor selisih antara tetangga minoritas terpilih (x_j) dan instans referensi itu sendiri. Vektor selisih ini diskalakan oleh sebuah skalar acak (λ) yang terdistribusi

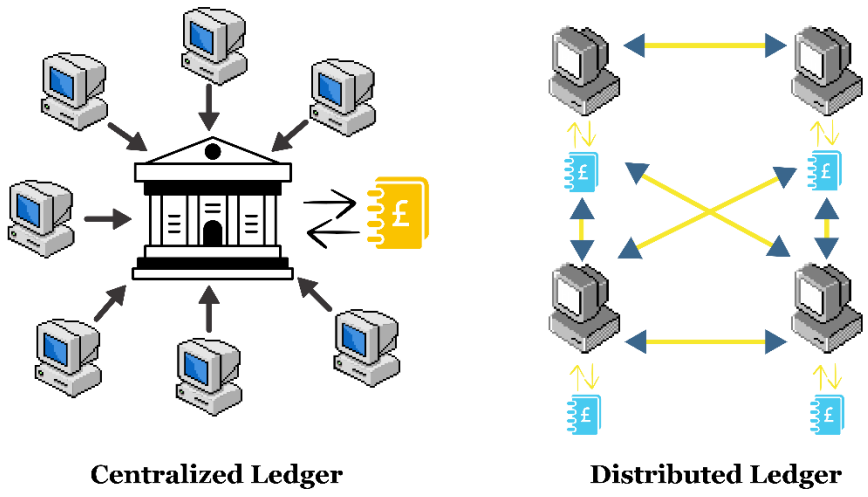
seragam pada interval [0, 1]. Mekanisme ini memastikan bahwa instans sintetis yang dihasilkan berada pada *hyper-line* segment yang menghubungkan dua instans minoritas yang ada, sehingga mempertahankan karakteristik distribusi data lokal dari kelas minoritas tersebut.

2.2.9 Distributed Ledger Technology (DLT)

Distributed Ledger Technology (DLT) mewakili transformasi paradigmatik dalam pencatatan digital, yang secara fundamental mengubah manajemen data dengan memungkinkan pemrosesan transaksi yang terdesentralisasi, tidak dapat diubah, dan transparan di seluruh jaringan terdistribusi (IEEE, 2023).

Berbeda dengan desain database konvensional yang bergantung pada otoritas pusat, DLT berjalan pada *topologi peer-to-peer* di mana setiap peserta menyimpan salinan buku besar yang tersinkronisasi dan identik, dengan semua perubahan divalidasi dan diimplementasikan melalui prosedur konsensus. Perbedaan arsitektur fundamental antara model terpusat dan terdistribusi ini diilustrasikan pada Gambar 2.4. Arsitektur struktural ini meniadakan kebutuhan akan perantara tradisional, sehingga meningkatkan integritas data, keamanan, dan kepercayaan di antara para anggota dalam sebuah jaringan.

Meskipun umumnya digunakan secara bergantian dengan blockchain, DLT adalah kategori teknologi yang lebih besar; blockchain adalah implementasi DLT spesifik yang mengatur data dalam rantai blok yang terurut secara kronologis dan terhubung secara kriptografis (Jia et al., 2023).



Gambar 2.4 Perbandingan Arsitektur *Ledger* Terpusat dan Terdistribusi.

Topologi DLT lainnya, seperti *Directed Acyclic Graphs* (DAGs), ada untuk menawarkan jawaban potensial terhadap batas penskalaan yang melekat pada berbagai implementasi blockchain (Moura et al., 2024). Kerangka kerja operasional jaringan DLT ditentukan lebih lanjut oleh tipologi kontrol aksesnya - membedakan antara sistem tanpa izin (publik) yang mengizinkan partisipasi tanpa batas dan jaringan berizin (privat) dengan akses terbatas - dan mekanisme konsensus yang memastikan kesepakatan di seluruh jaringan (Bonnet & Teuteberg, 2023).

Protokol konsensus yang menonjol termasuk *Proof of Work* (PoW) yang intensif secara komputasi, *Proof of Stake* (PoS) berbasis modal, dan protokol pengiriman pesan seperti *Practical Byzantine Fault Tolerance* (PBFT), yang masing-masing menyajikan pertukaran yang berbeda antara keamanan, *throughput* transaksi, dan efisiensi energi (Gorbunova et al., 2022). Aplikasi DLT sangat luas, mencakup layanan keuangan melalui tokenisasi aset, manajemen rantai pasokan melalui peningkatan pemantauan sumber, dan industri energi dengan mendukung perdagangan energi *peer-to-peer* dan mengelola sumber daya energi yang terdistribusi (IEEE, 2023). Terlepas dari manfaatnya yang luar biasa dalam hal disintermediasi, transparansi, dan keamanan, adopsi DLT secara umum menghadapi rintangan yang cukup besar, termasuk skalabilitas teknis, ambiguitas legislatif, dan komplikasi dalam tata kelola serta interaksi dengan sistem yang lebih tua.

2.2.10 Smart Contracts

Smart contract merupakan protokol komputasi otonom yang dapat berjalan sendiri yang memfasilitasi, memverifikasi, dan menegakkan negosiasi atau eksekusi perjanjian digital tanpa campur tangan perantara (Szabo, 1997). Disebarkan pada buku besar terdistribusi berbasis *blockchain*, konstruk yang dapat diprogram ini beroperasi secara deterministik melalui logika berbasis kondisi, menjalankan operasi yang telah ditentukan ketika kriteria tertentu terpenuhi (W. Zou et al., 2021). Sifatnya yang tidak dapat diubah, terdesentralisasi, dan transparan pada dasarnya mengubah kerangka kerja kontrak di berbagai domain, terutama dalam layanan keuangan, manajemen rantai pasokan, dan sistem energi.

Arsitektur operasional *smart contract* mengintegrasikan tiga properti *blockchain* yang penting: eksekusi yang terdesentralisasi, mekanisme keamanan kriptografi, dan proses komputasi deterministik (Goudarzi et al., 2022). Tidak seperti pengaturan kontrak

konvensional yang membutuhkan penegakan pihak ketiga, smart contract dieksekusi secara otonom di seluruh jaringan terdistribusi melalui protokol validasi konsensus. Kekekalannya setelah penerapan mencegah modifikasi sepihak, memastikan integritas transaksional dan membangun jalur eksekusi yang dapat diaudit (Alshahrani et al., 2023). Kerangka kerja arsitektur ini memungkinkan respons otomatis terhadap kondisi yang telah ditentukan misalnya, menerapkan hukuman untuk konsumsi listrik yang curang melalui pembatasan akun segera setelah deteksi anomali (Sayeed et al., 2020).

Dalam lingkungan smart grid yang mendukung IoT, implementasi *smart contract* memfasilitasi berbagai fungsi penting: (1) mekanisme respons anomali otomatis yang memicu tindakan yang telah ditentukan sebelumnya setelah mendeteksi pola konsumsi yang tidak teratur (Q. Yang & Wang, 2021); (2) platform perdagangan energi *peer-to-peer* yang memungkinkan transaksi langsung antara konsumen dan produsen melalui struktur harga yang diterapkan secara algoritmik (W. Liu & P, 2024); dan (3) sistem perekaman data tahan gangguan yang menyimpan catatan konsumsi yang tidak dapat diubah untuk kepatuhan terhadap peraturan dan tujuan audit (C. Hu et al., 2024). Implementasi fungsi-fungsi ini ditingkatkan melalui bahasa pemrograman *Turing-Completeness* Ethereum (Solidity), sementara kendala skalabilitas diatasi melalui solusi Layer-2 seperti Polygon, yang secara signifikan mengurangi biaya transaksi dan meningkatkan kapasitas keluaran (Labs, 2023).

Terlepas dari potensi transformatifnya, implementasi *smart contract* menghadapi tantangan besar, termasuk: kerentanan kerentanan kode, yang dibuktikan dengan pelanggaran keamanan historis yang mengakibatkan kerugian finansial yang signifikan, ambiguitas yurisdiksi terkait keberlakuan hukum perjanjian berbasis kode di berbagai kerangka kerja peraturan yang berbeda dan keterbatasan skalabilitas yang dimanifestasikan melalui biaya transaksi yang tinggi selama periode kemacetan jaringan (Richter Vidal et al., 2024). Strategi mitigasi meliputi metodologi verifikasi formal untuk deteksi kerentanan, kerangka kerja kontrak hibrida legal-smart yang mengintegrasikan mekanisme penyelesaian sengketa tradisional, dan pengoptimalan arsitektur yang meningkatkan efisiensi komputasi (Capocasale & Perboli, 2022).

2.2.11 Consensus Mechanisms

Mekanisme konsensus merupakan protokol dasar yang memungkinkan jaringan terdistribusi untuk mencapai kesepakatan mengenai keabsahan transaksi dan status buku

besar tanpa campur tangan otoritas terpusat (Nakamoto, 2008). Kerangka kerja algoritmik ini memastikan *Byzantine Fault Tolerance* mempertahankan integritas sistem meskipun ada node yang berbahaya atau rusak-sambil menetapkan finalitas transaksi dan mencegah serangan pembelanjaan ganda (Lamport et al., 1982). Pemilihan protokol konsensus yang tepat secara signifikan berdampak pada parameter kinerja jaringan, termasuk kapasitas throughput, latensi, konsumsi energi, dan tingkat jaminan keamanan.

Taksonomi konsensus saat ini mencakup beberapa kategori utama, masing-masing menawarkan karakteristik operasional yang berbeda dan pertukaran kinerja. Mekanisme *Proof of Work* (PoW) membutuhkan penyelesaian teka-teki komputasi untuk validasi blok, memberikan keamanan yang kuat melalui komitmen sumber daya tetapi membebankan konsumsi energi yang besar dan keterbatasan throughput (Nakamoto, 2008). Protokol *Proof of Stake* (PoS) memilih validator secara proporsional dengan taruhan ekonomi mereka, mengurangi kebutuhan energi sekitar 99% sambil mempertahankan jaminan keamanan yang sebanding melalui disinsentif ekonomi untuk perilaku jahat (Buterin, 2020). Implementasi *Practical Byzantine Fault Tolerance* (PBFT) menawarkan throughput transaksi yang tinggi dan latensi yang rendah tetapi membutuhkan set validator yang diketahui, sehingga membatasi potensi desentralisasi. *Delegated Proof of Stake* (DPoS) bertujuan untuk meningkatkan skalabilitas dengan meminta peserta jaringan memilih sejumlah validator terbatas, atau "*delegator*" untuk menghasilkan dan memvalidasi blok. Meskipun ini meningkatkan throughput transaksi, hal itu memperkenalkan vektor sentralisasi yang terus menjadi subjek analisis dalam studi tata kelola DLT terbaru (Y. Zou et al., 2020).

Dalam sistem manajemen energi, pemilihan mekanisme konsensus memerlukan pertimbangan cermat terhadap persyaratan aplikasi tertentu. Protokol *throughput* tinggi, seperti varian *Proof of Stake* (PoS), dapat memfasilitasi deteksi anomali waktu nyata dalam pemantauan konsumsi listrik, memungkinkan identifikasi aktivitas penipuan yang cepat dengan kepastian yang sesuai untuk infrastruktur kritis. Demikian pula, implementasi perdagangan energi mikrogrid mendapat manfaat dari validasi transaksi berbiaya rendah, memungkinkan pertukaran energi antar teman sebaya tanpa gesekan tanpa biaya *overhead* yang mahal (Kaur et al., 2021). Menerapkan mekanisme ini membutuhkan penanganan trade-off inheren antara desentralisasi, keamanan, dan skalabilitas, yang sering disebut "*blockchain trilemma*" (Principato et al., n.d.). Misalnya,

implementasi *Proof of Work* memberikan desentralisasi yang signifikan tetapi dengan throughput yang terbatas, sementara protokol konsensus berizin seperti *Practical Byzantine Fault Tolerance* (PBFT) dapat mencapai ribuan transaksi per detik tetapi dengan biaya kepercayaan yang terpusat (Wei et al., 2020). Penyelarasan peraturan menghadirkan pertimbangan tambahan, karena implementasi yang diizinkan menggunakan varian PBFT sering dianggap lebih kompatibel dengan kerangka peraturan yang ada di sektor-sektor seperti energi (Xiao et al., 2020).

Potensi teoritis dari *blockchain* sebagai sebuah *database* yang tidak dapat diubah (*immutable*) hanya dapat diwujudkan melalui mekanisme konsensusnya-protokol yang mengontrol bagaimana *node* terdistribusi menyetujui keaslian transaksi. Akan tetapi, pemilihan metode bukanlah sebuah masalah yang sepele; ini merupakan sebuah keputusan desain utama yang secara langsung mempengaruhi skalabilitas sistem, biaya, dan yang paling penting adalah kemampuannya untuk menjadi sebuah fondasi kepercayaan yang benar-benar terdesentralisasi. Pemeriksaan terhadap model konsensus saat ini menemukan bahwa banyak model yang pada dasarnya tidak sesuai dengan tuntutan sistem verifikasi IoT berskala besar dan *real-time*.

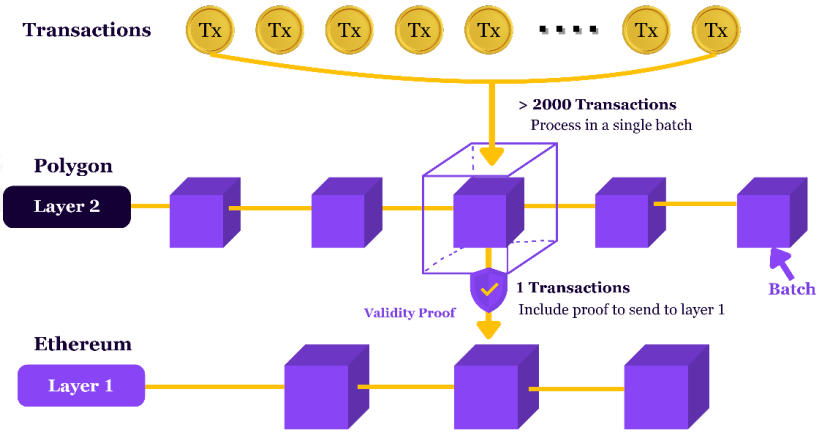
Model original dari *Proof-of-Work* (PoW) misalnya, jelas tidak sesuai. Ketergantungannya pada pemrosesan yang ekstensif menjamin keamanan yang kuat tetapi dengan biaya konsumsi energi yang sangat tinggi dan hasil transaksi yang sangat rendah (Nakamoto, 2008). Untuk sebuah sistem yang dimaksudkan untuk merekam kejadian anomali yang mungkin sering terjadi dari ribuan perangkat, PoW tidak dapat digunakan secara ekonomi dan lingkungan. Sebaliknya, arsitektur dengan throughput tinggi yang diijinkan seperti *Practical Byzantine Fault Tolerance* (PBFT) memiliki masalah yang terpisah dan lebih mendasar. Meskipun cepat dan efisien, mereka membutuhkan kumpulan validator yang diketahui dan terbatas (Wei et al., 2020). Solusi ini memperkenalkan kembali semacam sentralisasi, sehingga gagal untuk mengatasi masalah dasar asimetri kepercayaan di antara para pemangku kepentingan. Buku besar yang diautentikasi hanya oleh *node* perusahaan utilitas, misalnya, tidak menawarkan jaminan yang lebih independen kepada konsumen dibandingkan dengan basis data terpusat tradisional.

Hal ini mengarah pada pemilihan strategis jaringan *Proof-of-Stake* (PoS) publik, seperti *sidechain* Polygon, sebagai fondasi yang paling tepat untuk TDAVF. Pendekatan

ini langsung menjawab "*blockchain trilemma*" dengan memberikan keseimbangan pragmatis antara keamanan, skalabilitas, dan desentralisasi. Paradigma PoS memungkinkan *throughput* yang jauh lebih tinggi dan biaya transaksi yang jauh lebih rendah daripada PoW, sehingga praktis secara ekonomi untuk mencatat "*verified fact*" dalam skala besar (Altarawneh et al., 2020). Yang terpenting, dengan menggunakan jaringan publik dan tanpa izin, ini menghindari masalah asimetri kepercayaan dari rantai pribadi. Keamanan dijaga oleh kelompok validator yang luas, insentif ekonomi, dan tersebar secara global, membentuk lapisan verifikasi yang netral dan dapat diaudit yang dapat dipercaya oleh semua pihak. Pilihan yang dipilih ini secara langsung mendukung konsep TDAVF tentang Skalabilitas dan Yayasan Kepercayaan yang Terdesentralisasi.

2.2.12 Polygon Network

Polygon Network mewakili infrastruktur penskalaan Layer-2 yang canggih yang dirancang untuk meningkatkan fungsionalitas Ethereum dengan mengatasi keterbatasan mendasar dalam throughput, latensi transaksi, dan efisiensi biaya (Labs, 2023). Awalnya dikembangkan sebagai *Matic Network*, protokol ini mengimplementasikan kerangka kerja arsitektur berlapis-lapis untuk membangun jaringan *blockchain* yang saling terhubung, yang secara efektif mensintesis jaminan keamanan Ethereum yang kuat dengan karakteristik skalabilitas yang ditingkatkan dari rantai independen. Konfigurasi arsitektur ini membuat Polygon sangat cocok untuk penerapan *Internet of Things* (IoT) di lingkungan *smart grid* yang membutuhkan pemrosesan transaksi frekuensi tinggi dan biaya operasional yang minimal (Mollah et al., 2021).



Gambar 2.5 Cara kerja *polygon layer*-2.

Arsitektur teknis Polygon terdiri dari tiga komponen utama yang secara kolektif membentuk profil kinerjanya. Protokol ini mengimplementasikan lingkungan pengembangan modular yang mendukung implementasi *blockchain* yang beragam (jaringan mandiri, rantai aman, dan solusi Layer-2 (seperti yang ditunjukkan pada Gambar 2.5), memungkinkan pembuatan jaringan khusus domain yang dioptimalkan untuk aplikasi energi. Selain itu, Polygon menggunakan sistem pemeriksaan *Proof-of-Stake* yang secara berkala melakukan batch transaksi ke mainnet Ethereum, mencapai sekitar 65.000 transaksi per detik - peningkatan yang signifikan dibandingkan dengan kapasitas asli Ethereum sebesar ~15 TPS - dengan tetap mempertahankan prinsip-prinsip desentralisasi (Gangwal et al., 2023). Selain itu, jaringan ini menerapkan model keamanan ganda yang menggabungkan keamanan konsensus Ethereum untuk finalitas transaksi dengan jaringan validator khusus Polygon, memberikan fleksibilitas implementasi di berbagai penerapan smart grid mulai dari platform perdagangan energi publik hingga jaringan utilitas berizin (Rebello et al., 2024).

Keunggulan teknis Polygon memungkinkan beberapa aplikasi transformatif dalam sistem energi. Waktu konfirmasi blok sub-detik protokol dan biaya transaksi minimal (sekitar $ 0,01 per transaksi) memfasilitasi pemantauan terus menerus terhadap pola konsumsi listrik di seluruh jaringan sensor IoT yang terdistribusi (Aliaga, 2025). Kemampuan ini memungkinkan deteksi anomali secara real-time dalam konsumsi listrik, dengan setiap titik data atau log anomali yang direkam secara permanen dan dapat diakses untuk verifikasi pemangku kepentingan. Efisiensi transaksi sistem ini mendukung fungsionalitas pembayaran mikro yang penting untuk perdagangan energi *peer-to-peer* antara konsumen, dengan implementasi yang telah terbukti menangani lebih dari satu juta transaksi harian (Neiheiser et al., 2023). Selain itu, kompatibilitas *Ethereum Virtual Machine* (EVM) Polygon memungkinkan integrasi tanpa batas dengan smart contract berbasis Solidity yang sudah ada untuk penagihan otomatis, audit energi, dan kepatuhan terhadap peraturan di jaringan distribusi listrik (M. Li et al., 2020). Interoperabilitas ini merupakan keuntungan penting untuk membangun mekanisme tata kelola energi yang aman, terdesentralisasi, dan otomatis melalui smart contract dan aplikasi terdesentralisasi.

Ketika dievaluasi terhadap infrastruktur blockchain alternatif untuk aplikasi energi, Polygon menunjukkan keunggulan kinerja yang substansial. Kapasitas transaksinya sebesar 65.000 TPS secara signifikan melebihi mainnet Ethereum (15 TPS) dan solusi

Layer-2 yang bersaing seperti Arbitrum (4.000 TPS). Biaya transaksi rata-rata di Polygon tetap sekitar $ 0,01, dibandingkan dengan $ 5- $ 50 di Ethereum mainnet dan $ 0,10- $ 0,30 pada platform Layer-2 yang bersaing. Penyelesaian transaksi terjadi dalam 2-3 detik di Polygon, dibandingkan dengan 5-6 menit di Ethereum mainnet dan 10-15 menit pada solusi Layer-2 lainnya. Selain itu, konsumsi energi Polygon mewakili sekitar 0,01% dari sistem *Proof-of-Work*, dibandingkan dengan kebutuhan energi Ethereum sebelum Penggabungan dan 0,1%-1% untuk solusi Layer-2 yang bersaing (William Cong et al., 2023).

Terlepas dari keunggulan teknisnya, implementasi Polygon menghadirkan beberapa tantangan. Meskipun menawarkan skalabilitas yang lebih baik dibandingkan dengan mainnet Ethereum, kumpulan validator Polygon yang relatif lebih kecil (sekitar 100 validator versus 300.000+ validator Ethereum) memperkenalkan potensi kerentanan sentralisasi (William Cong et al., 2023). Beberapa insiden keamanan terkenal yang melibatkan jembatan *blockchain* menekankan potensi kerentanan dalam transfer data energi lintas rantai antara Polygon dan jaringan lain (Gupta et al., 2024). Selain itu, arsitektur *hybrid* Polygon menciptakan pertimbangan yurisdiksi yang kompleks untuk kepatuhan pasar energi, terutama terkait kedaulatan data dan otoritas validasi transaksi (Laayati et al., 2022).

Peta jalan teknologi Polygon mengatasi keterbatasan saat ini melalui beberapa inisiatif strategis. Penerapan bukti kriptografi tanpa pengetahuan bertujuan untuk meningkatkan pelestarian privasi untuk data konsumsi energi yang sensitif dengan tetap menjaga skalabilitas komputasi. Pengembangan "*Supernets*" konfigurasi blockchain khusus yang dirancang untuk koperasi energi atau penyedia utilitas tertentu memungkinkan pengoptimalan khusus domain (Labs, 2023).

2.2.13 Edge Computing

Edge Computing mewakili paradigma komputasi terdistribusi yang menempatkan sumber daya komputasi di dekat sumber penghasil data, sehingga meminimalkan latensi dan pemanfaatan *bandwidth* sekaligus meningkatkan kemampuan pemrosesan waktu nyata. Pendekatan arsitektur ini secara fundamental mengubah infrastruktur tradisional yang berpusat pada *cloud* dengan memfasilitasi pemrosesan data yang terlokalisasi di pinggiran jaringan-khususnya di gardu induk, trafo, dan meteran pintar-daripada mentransmisikan semua data ke tempat penyimpanan yang terpusat (Qiu et al., 2020).

Dalam konteks sistem *smart grid*, komputasi *edge* diwujudkan melalui arsitektur hirarkis yang terdiri dari perangkat *edge* (smart meter, Unit Pengukuran *Phasor*), *gateway* perantara, dan *server edge* yang terlokalisasi. Infrastruktur ini memungkinkan fungsi-fungsi penting termasuk deteksi anomali waktu nyata, pengambilan keputusan secara otonom, dan integrasi *blockchain* yang efisien (Mehmood et al., 2021). Dengan menganalisis pola konsumsi listrik secara lokal melalui model *machine learning* yang dioptimalkan (misalnya, LightGBM yang dikuantifikasi, CNN), sistem *edge* dapat mengidentifikasi ketidakberesan seperti konsumsi yang tidak sah dan kerusakan peralatan dalam hitungan milidetik, yang secara substansial mengurangi latensi respons dibandingkan dengan arsitektur yang bergantung pada *cloud* (S. Chen et al., 2019).

Integrasi komputasi *edge* dengan teknologi *blockchain* menghasilkan aplikasi yang sangat menjanjikan dalam sistem energi pintar. *Edge node* melakukan pra-validasi data sensor sebelum dikirim ke buku besar terdistribusi seperti Polygon, sehingga meminimalkan transaksi yang berlebihan dan biaya komputasi yang terkait dengan tetap menjaga integritas data (Labs, 2023). Pendekatan hibrida ini memfasilitasi eksekusi *smart contract* yang dipercepat untuk operasi yang sangat penting sekaligus mengurangi biaya transaksi *blockchain* melalui pengiriman data secara selektif (Aliaga, 2025).

Terlepas dari kelebihannya, implementasi *edge computing* menghadapi beberapa kendala, termasuk sumber daya komputasi yang terbatas sehingga memerlukan optimasi model, kerentanan keamanan yang meningkat karena aksesibilitas fisik, dan tantangan interoperabilitas di seluruh ekosistem perangkat yang heterogen (R. Liu et al., 2022). Arah penelitian yang muncul mengatasi keterbatasan ini melalui pendekatan pembelajaran federasi yang memungkinkan deteksi anomali kolaboratif tanpa transfer data mentah, *hybrid edge-blockchain* yang menggabungkan bukti pengetahuan nol untuk pengiriman data terverifikasi, dan integrasi dengan infrastruktur telekomunikasi canggih untuk mendukung komunikasi latensi rendah yang sangat andal untuk operasi jaringan yang sangat penting (J. Li et al., 2022).

2.3 Kelemahan dan Gap Penelitian Terdahulu

Kesenjangan paling mendasar dan signifikan adalah tidak adanya kerangka kerja arsitektur yang berpusat pada verifikasi. Penelitian yang ada membahas deteksi anomali sebagai masalah pengenalan pola yang berakhir dengan keluaran deteksi, tanpa memeriksa bagaimana keluaran tersebut dapat diubah menjadi bukti yang dipercaya

secara universal seperti yang ditunjukkan pada Tabel 2.1.

Studi literatur menemukan bahwa tidak ada paradigma yang ada yang secara sistematis membahas transisi hasil deteksi menjadi fakta terverifikasi yang dapat diperiksa secara independen oleh banyak pihak. Kesenjangan konseptual ini mendorong ketergantungan terus-menerus pada otoritas pusat untuk penyelesaian konflik dan validasi bukti.

Kesenjangan terbesar yang teridentifikasi adalah tidak adanya kerangka kerja lengkap yang menghubungkan *edge computing* dengan kemampuan verifikasi terdesentralisasi. *Trust-Based Decentralized Anomaly Verification Framework* (TDAVF) mengatasi kekurangan ini dengan memperkenalkan perubahan paradigma dari pendekatan yang berpusat pada deteksi menjadi pendekatan yang berpusat pada verifikasi. Secara spesifik, TDAVF mengatasi kekurangan yang disorot melalui.

1. ***Architectural Integration***. Campuran sistematis pengumpulan data *real-time* IoT, peningkatan LightGBM untuk deteksi anomali berbasis *edge* yang efisien, dan integrasi blockchain untuk menghasilkan bukti yang tidak dapat diubah.
2. ***Edge-Optimized Processing***. Optimasi hiperparameter LightGBM untuk kinerja *real-time* pada perangkat dengan sumber daya terbatas sambil mempertahankan akurasi deteksi.
3. ***Decentralized Trust Foundation***. Penghapusan ketergantungan pada otoritas pusat melalui produksi bukti yang dapat diverifikasi secara kriptografis dan dapat diaudit oleh semua pemangku kepentingan.
4. ***Scalability Framework***. Desain hibrida *edge-blockchain* yang menghilangkan ketergantungan pada cloud sambil mempertahankan skalabilitas sistem untuk penyebaran infrastruktur energi skala besar.

Pendekatan ini tidak hanya meningkatkan akurasi deteksi tetapi juga secara fundamental mengubah manajemen anomali energi dari otoritas berbasis kepercayaan menjadi transparansi yang diverifikasi secara kriptografis, mengatasi penyebab utama tantangan yang sedang berlangsung dalam ekosistem energi.

Tabel 2.1 Studi tentang Penerapan ML, Blockchain dan IoT dalam Sektor Energi

Aspek	Penulis	Metode yang sudah ada	Kontribusi Utama	Keterbatasan	Novelty Penelitian
Data Architecture	Zhuang et al. (Zhuang et al., 2024)	GCN dengan *residual learning*	Meningkatkan akurasi deteksi menggunakan model generatif	Tidak memiliki data waktu nyata IoT, mahal secara komputasi, bergantung pada pemrosesan terpusat.	TDAVF mengatasinya dengan arsitektur terdesentralisasi yang memvalidasi deteksi di *edge* dan mencatatnya sebagai bukti *immutable* di *ledger*, memastikan jejak audit yang utuh.
	Cai et al. (Cai et al., 2023)	*Random Forest* + SVDD	Meningkatkan akurasi deteksi anomali menggunakan pembelajaran statistik	Tidak ada integrasi IoT, tidak memiliki blockchain, *false positive* yang tinggi.	
	Zhang et al. (W. Zhang et al., 2020)	*Feature engineering* dengan *Density-Based Clustering*	Meningkatkan deteksi anomali tanpa data berlabel	Kemampuan terbatas untuk mendeteksi jenis anomali tertentu, kurangnya implementasi *real-time*.	
Anomaly Detection	Wu et al. (Wu & Wu, 2024)	Model Hybrid CNN-BiLSTM dengan *self-attention*	Meningkatkan akurasi prediksi	Latensi tinggi; tidak cocok untuk perangkat *edge.*	LightGBM pada perangkat *edge* dengan *hyperparameter* yang dioptimalkan untuk perangkat dengan sumber daya rendah.
	Bian et al. (Bian et al., 2021)	PSO untuk mengoptimalkan model LSTM-Attention	Meningkatkan deteksi anomali secara dinamis	*High computational cost.*	

Aspek	Penulis	Metode yang sudah ada	Kontribusi Utama	Keterbatasan	Novelty Penelitian
	Taruna et al. (Taruna et al., 2025)	Model ML yang mengoptimalkan *Target Operation* (TO)	Mencapai akurasi 0,89 menggunakan *Random Forest* & KNN	Tidak memiliki integrasi waktu nyata IoT.	
Privacy & Data Sharing	Zhao et al. (Z. Zhao et al., 2024)	Menggunakan enkripsi homomorfik untuk analisis *smart meter* yang aman	Meningkatkan perlindungan privasi	Tidak ada *blockchain*, tidak memiliki integrasi *real-time* IoT.	Tokenisasi Web 3.0 untuk berbagi data milik pengguna dengan izin granular.
	Zhao et al. (Z. Zhao et al., 2023)	Deteksi anomali aman berbasis blockchain	Memastikan penyimpanan data yang transparan	Pengguna tidak dapat mengontrol akses data; berbagi tidak jelas.	
Otomatisasi	Civelek et al. (Civelek et al., 2024)	Menggunakan FWHT dengan ML untuk deteksi pencurian	Menyediakan deteksi pencurian geospasial	Respons yang tertunda terhadap pencurian; kesalahan manusia.	Otomatisasi *smart contract* untuk penyesuaian dan peringatan penagihan instan.
	Nabil et al. (Nabil et al., 2019)	Model Hybrid *ensemble learning*	Meningkatkan prediksi jangka pendek	Berfokus pada peramalan, tidak memiliki notifikasi IoT dan peringatan waktu nyata.	
Skalabilitas	Din et al. (Din et al., 2024)	Menggunakan *blockchain* untuk transaksi energi P2P yang aman	Memastikan transaksi yang tahan gangguan	Biaya operasional yang tinggi; latensi dalam jaringan yang besar.	Gabungan antata *edge-blockchain* untuk pemrosesan terdistribusi, mengurangi ketergantungan *cloud*.