UNIVERSITAS ESA UNGGUL

Perbandingan Algoritma *XGBoost-SMOTE* dan *Random Forest-SMOTE* untuk Deteksi *Fraud* Transaksi Kartu Kredit Berdasarkan Standar POJK 2022

PROPOSAL TUGAS AKHIR

Diajukan untuk Memenuhi Sebagian dari Syarat-syarat Tugas Akhir

Nama : Putri Agisnadiansyah

NIM : 20220801330

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS ESA UNGGUL

2025

KATA PENGANTAR

Puji syukur peneliti panjatkan ke hadirat Allah SWT atas limpahan rahmat, karunia, serta kesempatan yang diberikan sehingga peneliti dapat menyelesaikan penyusunan proposal Tugas Akhir yang berjudul "Perbandingan Algoritma *XGBoost–SMOTE* dan *Random Forest–SMOTE* untuk Deteksi *Fraud* Transaksi Kartu Kredit Berdasarkan Standar POJK 2022" dengan baik dan lancar.

Penyusunan proposal ini merupakan salah satu syarat untuk menyelesaikan studi Strata-1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Esa Unggul. Dalam proses penyusunannya, peneliti memperoleh banyak dukungan, bimbingan, dan arahan dari berbagai pihak. Oleh karena itu, peneliti ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Kedua orang tua dan keluarga tercinta, atas kasih sayang, doa, dukungan moral maupun material, serta semangat yang tidak pernah putus dalam setiap langkah peneliti selama menjalani perkuliahan hingga penyusunan Tugas Akhir ini.
2. Bapak Dr. Ir. Arief Kusuma A.P., S.T., MBA., IPU, ASEAN Eng., selaku Rektor Universitas Esa Unggul, atas fasilitas dan kesempatan yang diberikan kepada mahasiswa untuk terus berkembang dalam bidang akademik maupun penelitian.
3. Bapak Dr. Gerry Firmansyah, S.T., M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas Esa Unggul, atas dukungan dan arahannya dalam menciptakan lingkungan akademik yang kondusif dan inovatif.
4. Ibu Diah Aryani, S.T., M.Kom., selaku Wakil Dekan Fakultas Ilmu Komputer Universitas Esa Unggul, atas perhatian dan dorongan yang diberikan kepada seluruh mahasiswa dalam mencapai keberhasilan akademik.
5. Ibu Dr. Riya Widayanti, S.Kom., MMSI., selaku Kepala Program Studi Teknik Informatika Universitas Esa Unggul, atas motivasi dan kebijakan yang mendukung kelancaran proses penyusunan Tugas Akhir ini.
6. Bapak Muhamad Bahrul Ulum, S.Kom., M.Kom., selaku Dosen Pembimbing Akademik, atas arahan dan pendampingan beliau selama peneliti menempuh studi di Universitas Esa Unggul, yang memberikan banyak pembelajaran dan pengalaman berharga.

7. Bapak Jefry Sunupurwa Asri, S.Kom., M.Kom., selaku Dosen Pembimbing Tugas Akhir, atas waktu, bimbingan, serta saran konstruktif yang diberikan dengan penuh kesabaran hingga proposal ini dapat terselesaikan dengan baik.
8. Teman-teman seperjuangan di Program Studi Teknik Informatika, yang telah memberikan dukungan, semangat, dan kebersamaan selama proses perkuliahan serta penyusunan Tugas Akhir ini.

Peneliti menyadari bahwa proposal ini masih memiliki keterbatasan dan kekurangan, oleh karena itu kritik dan saran yang membangun sangat diharapkan demi perbaikan di masa yang akan datang.

Akhir kata, semoga proposal ini dapat memberikan manfaat dan menjadi salah satu kontribusi kecil dalam pengembangan ilmu pengetahuan di bidang *Artificial Intelligence* dan *Financial Technology* yang sesuai dengan prinsip tata kelola risiko berdasarkan POJK 2022.

Tangerang, 1 November 2025

Putri Agisnadiansyah
(20220801330)

ABSTRAK

Judul	:	Perbandingan Algoritma *XGBoost-SMOTE* dan *Random Forest-SMOTE* untuk Deteksi *Fraud* Transaksi Kartu Kredit Berdasarkan Standar POJK 2022
Nama	:	Putri Agisnadiansyah
Program Studi	:	Teknik Informatika

Peningkatan transaksi digital di sektor perbankan telah meningkatkan risiko *fraud* kartu kredit, yang menuntut penerapan sistem deteksi akurat dan sesuai dengan prinsip manajemen risiko sebagaimana diatur dalam POJK 2022. Penelitian ini bertujuan membandingkan kinerja dua algoritma *machine learning* berbasis *ensemble*, yaitu *XGBoost–SMOTE* dan *Random Forest–SMOTE*, dalam mendeteksi transaksi *fraudulent* pada *dataset* kartu kredit yang tidak seimbang. Metodologi penelitian meliputi *data preprocessing*, penerapan *Synthetic Minority Over-sampling Technique*, pelatihan model, evaluasi metrik (*Recall*, *F1-Score*, dan *AUC-ROC*), serta analisis keselarasan hasil dengan prinsip *security*, *transparency*, dan *IT risk management* yang diatur oleh POJK 2022. Hasil eksperimen menunjukkan bahwa model *XGBoost–SMOTE* menampilkan kinerja yang lebih unggul dalam mendeteksi transaksi *fraud*, terutama pada metrik *Recall* dan *F1-Score*, dibandingkan dengan *Random Forest–SMOTE*. Studi ini menegaskan bahwa pendekatan *machine learning* tidak hanya efektif secara teknis, tetapi juga dapat disesuaikan dengan regulasi nasional untuk mengembangkan sistem deteksi *fraud* yang transparan, aman, dan *compliant* terhadap POJK 2022.

Kata kunci: *Fraud Detection, XGBoost, Random Forest, SMOTE, POJK 2022, Machine Learning, credit card transactions.*

ABSTRACT

Title	:	*A Comparative Study of XGBoost–SMOTE and Random Forest–SMOTE Algorithms for Credit Card Fraud Detection Based on the Standards of POJK 2022*
Name	:	*Putri Agisnadiansyah*
Study Program	:	*Informatics Engineering*

The rapid growth of digital banking transactions has increased the risk of credit card fraud, emphasizing the need for accurate and regulation-compliant detection systems as required by POJK 2022. This study aims to compare the performance of two ensemble-based machine learning algorithms, XGBoost–SMOTE and Random Forest–SMOTE, in detecting fraudulent credit card transactions within an imbalanced dataset. The research stages include data preprocessing, application of the Synthetic Minority Over-sampling Technique (SMOTE), model training, evaluation using metrics such as Recall, F1-Score, and AUC-ROC, and an analysis of how the results align with POJK 2022 principles of security, transparency, and IT risk management. Experimental results demonstrate that the XGBoost–SMOTE model performs better in identifying fraudulent transactions, particularly in terms of Recall and F1-Score, compared to Random Forest–SMOTE. This study highlights that machine learning approaches are not only technically effective but can also be adapted to meet national regulatory standards, supporting the development of transparent, secure, and compliant fraud detection systems for the banking sector.

Keywords*: Fraud Detection, XGBoost, Random Forest, SMOTE, POJK 2022, Machine Learning*

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi finansial (*financial technology/fintech*) telah meningkatkan volume transaksi digital secara signifikan, terutama dalam penggunaan kartu kredit. Namun, peningkatan tersebut juga diiringi dengan peningkatan risiko penipuan (*fraud*), yang dapat menyebabkan kerugian finansial bagi nasabah dan lembaga perbankan.

Dalam konteks Indonesia, Peraturan Otoritas Jasa Keuangan (POJK) tahun 2022 mengenai penerapan teknologi informasi oleh bank umum menekankan pentingnya aspek keamanan, auditabilitas, dan transparansi pada sistem berbasis kecerdasan buatan (*artificial intelligence*). Oleh karena itu, sistem deteksi *fraud* berbasis *machine learning* tidak hanya harus akurat, tetapi juga harus mematuhi prinsip regulatif yang diatur oleh POJK tersebut.

Algoritma *machine learning* seperti *Extreme Gradient Boosting* (*XGBoost*) dan *Random Forest* telah terbukti efektif dalam mendeteksi pola transaksi yang mencurigakan. Namun, salah satu tantangan utama dalam penerapan model ini adalah ketidakseimbangan data (*imbalanced dataset*), di mana jumlah transaksi *fraud* jauh lebih sedikit dibandingkan transaksi normal. Kondisi ini membuat model cenderung bias terhadap kelas mayoritas (*non-fraud*).

Untuk mengatasi masalah tersebut, digunakan teknik *Synthetic Minority Oversampling Technique* (*SMOTE*) yang bertujuan menyeimbangkan distribusi data tanpa kehilangan karakteristik penting. Berdasarkan hasil *Systematic Literature Review* (*SLR*) dan analisis *Business Requirement Document* (*BRD*) yang telah dilakukan, ditemukan adanya *research gap* berupa:

1. Belum adanya studi yang secara spesifik membandingkan *XGBoost-SMOTE* dan *Random Forest-SMOTE* dalam konteks kepatuhan terhadap POJK 2022.
2. Kurangnya integrasi antara *performance metrics*, *explainability analysis*, dan *regulatory compliance* dalam satu kerangka penelitian komprehensif.

Penelitian ini bertujuan untuk mengisi kesenjangan tersebut dengan melakukan analisis komparatif terhadap kedua algoritma, serta menilai sejauh mana model yang dibangun dapat mendukung prinsip keamanan, transparansi, dan auditabilitas sebagaimana diamanatkan oleh regulasi POJK 2022.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah penelitian ini adalah sebagai berikut:

1. Bagaimana merancang sistem deteksi *fraud* kartu kredit menggunakan kombinasi *XGBoost–SMOTE* dan *Random Forest–SMOTE* yang efektif dan efisien?
2. Bagaimana perbandingan kinerja kedua algoritma tersebut berdasarkan metrik evaluasi seperti *Recall, F1-Score*, *AUC-ROC*, *dan False Positive Rate*?
3. Bagaimana keterkaitan hasil model dengan prinsip keamanan, transparansi, dan auditabilitas sesuai ketentuan POJK 2022 dalam konteks deteksi *fraud* kartu kredit?
4. Bagaimana penerapan dan relevansi POJK 2022 dapat mendukung penguatan sistem deteksi *fraud* kartu kredit di sektor perbankan Indonesia?

1.3 Tujuan Penelitian

Tujuan utama dari penelitian ini adalah sebagai berikut:

1. Membandingkan performa algoritma *XGBoost–SMOTE* dan *Random Forest–SMOTE* dalam mendeteksi transaksi kartu kredit yang bersifat *fraudulent*.
2. Menganalisis efektivitas teknik *SMOTE* dalam mengatasi permasalahan *imbalanced dataset* pada sistem deteksi *fraud.*
3. Mengevaluasi keterkaitan hasil model dengan prinsip keamanan, transparansi, dan auditabilitas sebagaimana diatur dalam POJK 2022, untuk memastikan kesesuaian penerapan *machine learning* dengan standar manajemen risiko teknologi informasi.
4. Mengkaji peran dan relevansi POJK 2022 dalam mendukung implementasi sistem deteksi *fraud* kartu kredit yang efektif, aman, dan sesuai dengan ketentuan regulasi perbankan nasional.

1.4 Manfaat Penelitian

1. Manfaat Akademik

- Memberikan kontribusi ilmiah terhadap pengembangan model *machine learning* untuk kasus data tidak seimbang (*imbalanced data*) di sektor finansial.

- Menghasilkan kerangka analisis yang menghubungkan kinerja teknis algoritma dengan aspek kepatuhan terhadap regulasi keuangan di Indonesia.

2. Manfaat Praktis

- Menyediakan rekomendasi berbasis bukti bagi lembaga keuangan dalam penerapan teknologi deteksi *fraud* yang *regulation-aware*.
- Memberikan wawasan bagi industri perbankan mengenai bagaimana regulasi POJK 2022 dapat memperkuat mekanisme pencegahan dan deteksi *fraud* kartu kredit.

3. Manfaat Pengembangan Sistem

- Menjadi dasar perancangan sistem deteksi *fraud* yang selaras dengan prinsip keamanan dan manajemen risiko sebagaimana ditetapkan oleh POJK 2022.

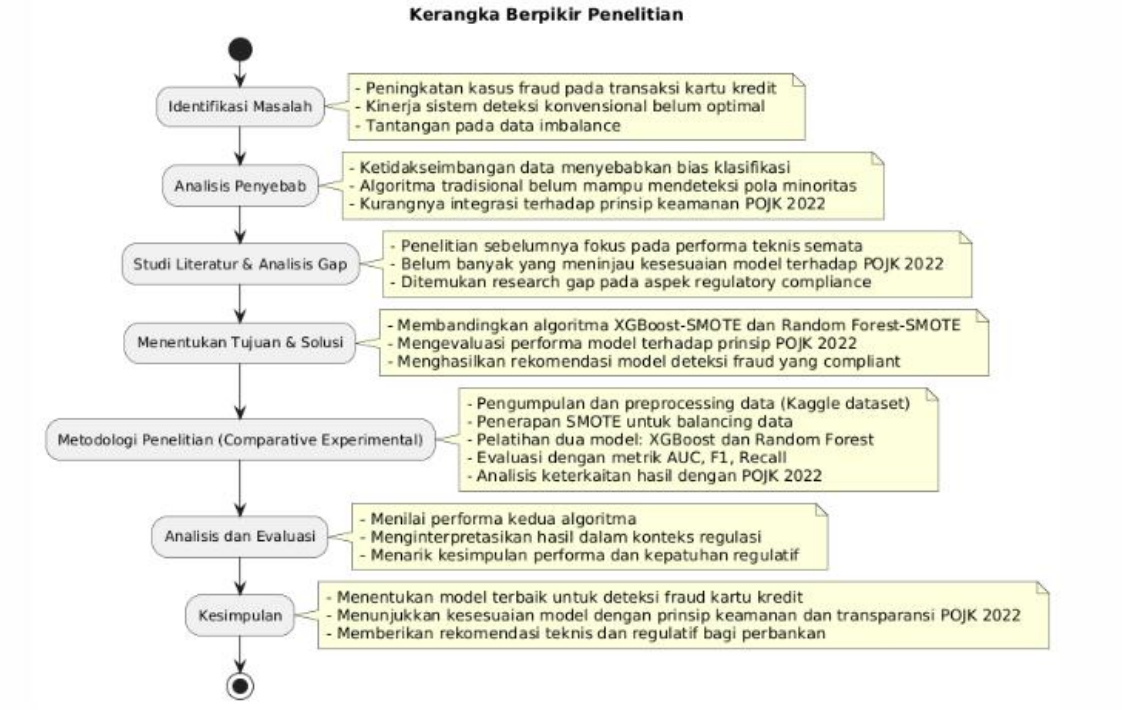
1.5 Batasan Penelitian

Agar penelitian ini tetap terarah dan fokus, maka batasan-batasan penelitian adalah sebagai berikut:

1. Penelitian ini hanya membandingkan dua algoritma utama, yaitu *XGBoost-SMOTE* dan *Random Forest-SMOTE*, tanpa melibatkan algoritma lain.
2. Evaluasi kinerja model difokuskan pada metrik *Recall*, *F1-Score*, dan *AUC-ROC*, sebagai indikator utama untuk data tidak seimbang.
3. Aspek kepatuhan regulatif hanya mengacu pada prinsip-prinsip yang tercantum dalam POJK 2022, tidak mencakup regulasi internasional seperti *GDPR* atau *Basel III*.
4. Implementasi sistem bersifat eksperimental dan dilakukan dalam lingkungan simulasi, bukan di sistem perbankan aktual.

1.6 Kerangka Berpikir Penelitian

Kerangka berpikir penelitian ini menggambarkan alur konseptual dari identifikasi masalah hingga analisis kesesuaian dengan regulasi POJK 2022.



Gambar 1.6.1 Kerangka Berpikir Penelitian

Penelitian ini diawali dengan identifikasi masalah berupa meningkatnya risiko *fraud* pada transaksi kartu kredit serta rendahnya efektivitas metode deteksi konvensional akibat ketidakseimbangan data. Tahap analisis penyebab menguraikan bahwa *imbalanced dataset* menyebabkan bias model dan lemahnya kemampuan sistem dalam mengenali pola *fraudulent transaction*.

Selanjutnya dilakukan studi literatur dan analisis gap, yang menemukan bahwa sebagian besar penelitian terdahulu hanya berfokus pada aspek performa teknis algoritma tanpa mempertimbangkan kesesuaian dengan prinsip-prinsip regulasi seperti *security*, *transparency*, dan *risk management* yang diatur dalam POJK 2022.

Tahap penentuan tujuan dan solusi merumuskan arah penelitian untuk melakukan perbandingan antara dua model *machine learning XGBoost–SMOTE* dan *Random Forest–SMOTE* dalam konteks deteksi *fraud* kartu kredit.

Pada tahap metodologi penelitian, dilakukan proses pengumpulan data, penerapan *SMOTE* untuk penyeimbangan data, pelatihan kedua model, dan evaluasi menggunakan metrik *AUC*, *Recall*, serta *F1-Score*.

Hasil dari proses ini kemudian dianalisis pada tahap analisis dan evaluasi, dengan mengaitkan performa model terhadap prinsip keamanan dan manajemen risiko POJK 2022. Akhirnya, pada tahap kesimpulan, ditarik hasil perbandingan dan disusun rekomendasi model

terbaik yang tidak hanya unggul secara teknis, tetapi juga *compliant* terhadap standar regulatif nasional.

1.7 Sistematika Penulisan

Adapun sistematika penulisan proposal tugas akhir ini disusun sebagai berikut:

BAB I Pendahuluan

Menjelaskan latar belakang, rumusan masalah, tujuan, manfaat, batasan penelitian, kerangka berpikir, dan sistematika penulisan.

BAB II Tinjauan Pustaka

Membahas teori-teori yang relevan, termasuk regulasi POJK 2022, algoritma yang digunakan (*XGBoost, Random Forest, SMOTE*), serta hasil penelitian terdahulu yang menjadi landasan penelitian ini.

BAB III Metodologi Penelitian

Menguraikan metode penelitian yang digunakan, termasuk desain eksperimen, dataset, tahapan implementasi, dan metode evaluasi kinerja model.

BAB IV Perancangan dan Implementasi

Menjelaskan rancangan arsitektur sistem, hasil eksperimen, dan analisis komparatif dari kedua algoritma.

BAB V Kesimpulan dan Saran

Berisi kesimpulan dari hasil penelitian serta saran untuk pengembangan penelitian selanjutnya.

BAB II

TINJAUAN PUSTAKA

2.1 *Fraud* Transaksi Kartu Kredit

Penipuan kartu kredit (*credit card fraud*) merupakan tindakan penyalahgunaan data kartu kredit secara ilegal untuk melakukan transaksi finansial tanpa izin pemilik yang sah. Berdasarkan laporan *Nilson Report* (2022), kerugian global akibat penipuan kartu kredit mencapai miliaran dolar setiap tahun dan terus meningkat seiring digitalisasi sistem pembayaran.

Dalam konteks perbankan Indonesia, peningkatan penggunaan *e-commerce* dan layanan perbankan digital membuat risiko *fraud* semakin tinggi. Otoritas Jasa Keuangan (OJK) melalui POJK 2022 menekankan pentingnya penerapan manajemen risiko teknologi informasi dan sistem keamanan yang mampu mendeteksi aktivitas transaksi yang mencurigakan secara *real-time*.

Fraud dalam transaksi kartu kredit umumnya memiliki karakteristik sebagai berikut:

1. *Highly Imbalanced Data* - Proporsi transaksi *fraud* jauh lebih kecil dari transaksi normal.
2. *Dynamic Behavior* - Pola *fraud* berubah-ubah seiring waktu karena pelaku terus beradaptasi.
3. *High Cost of False Negatives* - Kegagalan mendeteksi transaksi *fraud* dapat berdampak besar terhadap kerugian finansial dan reputasi bank.

Oleh karena itu, sistem deteksi *fraud* modern harus menggunakan pendekatan berbasis *machine learning* dengan kemampuan adaptif terhadap data baru dan ketidakseimbangan kelas.

2.2 Regulasi POJK 2022 dan Relevansinya terhadap Deteksi *Fraud*

POJK 2022 mengatur prinsip penerapan manajemen risiko teknologi informasi dan keamanan sistem perbankan digital. Regulasi ini menekankan tiga aspek utama yang berkaitan langsung dengan deteksi *fraud*, yaitu:

1. Keamanan Data dan Transaksi (*Data & Transaction Security*): Bank wajib menjamin integritas, kerahasiaan, dan ketersediaan data transaksi

keuangan. Sistem deteksi *fraud* harus mampu meminimalkan risiko kebocoran data dan penyalahgunaan transaksi.

2. Auditabilitas dan Akuntabilitas Sistem (*System Auditability*): Seluruh aktivitas sistem berbasis *AI*, termasuk *machine learning models*, harus dapat diaudit dan dilacak kembali proses keputusannya. Ini menjadi dasar penting bagi pengembangan sistem deteksi *fraud* yang transparan dan dapat diverifikasi.

3. Penerapan Manajemen Risiko Teknologi Informasi (*IT Risk Management*): POJK 2022 mewajibkan lembaga keuangan melakukan evaluasi risiko terhadap sistem berbasis *AI*, termasuk potensi *false alarm* pada sistem deteksi *fraud* yang dapat mempengaruhi kepercayaan nasabah.

Keterkaitan antara POJK 2022 dan deteksi *fraud* kartu kredit dapat dilihat dari bagaimana prinsip-prinsip tersebut mengarahkan sistem *AI* agar tidak hanya unggul secara performa teknis, tetapi juga sejalan dengan tata kelola dan kepatuhan (*regulatory compliance*) di sektor perbankan Indonesia.

2.3 *Machine Learning* untuk Deteksi *Fraud*

Deteksi *fraud* menggunakan *machine learning* berfungsi untuk mengidentifikasi pola anomali pada data transaksi. Pendekatan ini berbeda dari sistem berbasis aturan (*rule-based system*), karena mampu belajar dari data historis dan menyesuaikan terhadap perubahan perilaku pelaku *fraud*.

Secara umum, proses deteksi *fraud* berbasis *machine learning* terdiri dari beberapa tahap:

1. *Data Preprocessing* – Pembersihan data, normalisasi, dan reduksi dimensi.
2. *Data Balancing* – Penanganan ketidakseimbangan data dengan teknik seperti SMOTE.
3. *Model Training* – Penerapan algoritma seperti *XGBoost* dan *Random Forest*.
4. *Evaluation* – Pengukuran performa menggunakan metrik seperti *Recall, Precision, F1-Score*, dan *AUC-ROC*.

5. *Interpretability & Compliance Analysis* – Analisis hasil prediksi model dalam konteks regulatif (misalnya, kesesuaian dengan prinsip POJK).

Beberapa studi sebelumnya (Brown & Mues, 2012; Dal Pozzolo et al., 2015) menunjukkan bahwa model *ensemble* memiliki kinerja yang lebih baik dibanding model tunggal dalam kasus deteksi *fraud* karena kemampuannya menangani *noise* dan *overfitting*.

2.4 Algoritma yang Digunakan

2.4.1 *XGBoost* (*Extreme Gradient Boosting*)

*XGBoos*t adalah algoritma *ensemble learning* berbasis *gradient boosting* yang menggabungkan banyak pohon keputusan (*decision trees*) secara iteratif untuk meminimalkan kesalahan prediksi. Keunggulan *XGBoost* antara lain:

- Kemampuan menangani data berukuran besar dengan efisiensi tinggi.
- Mendukung penanganan *imbalanced dataset* melalui parameter *scale_pos_weight*.
- Dapat dioptimalkan menggunakan *cross-validation* untuk meningkatkan akurasi model.

Studi oleh Chen & Guestrin (2016) menunjukkan bahwa *XGBoost* sering unggul dalam kompetisi deteksi anomali karena kestabilan dan akurasi tinggi pada data kompleks.

2.4.2 *Random Forest*

Random Forest merupakan algoritma *bagging ensemble* yang menggabungkan banyak pohon keputusan secara acak untuk menghasilkan prediksi yang stabil. Kelebihan utamanya meliputi:

- *Robust* terhadap *overfitting* dan *noise*.
- Mudah diinterpretasikan melalui analisis *feature importance*.
- Cocok untuk data berukuran besar dan tidak terstruktur.

Menurut Breiman (2001), *Random Forest* cenderung memberikan hasil yang lebih stabil dibanding model *boosting*, meskipun kadang kalah dalam hal sensitivitas terhadap kelas minoritas.

2.5 Teknik Penanganan Data Tidak Seimbang

2.5.1 *SMOTE* (*Synthetic Minority Oversampling Technique*)

SMOTE (Chawla et al., 2002) merupakan metode *oversampling* yang menghasilkan sampel sintetik untuk kelas minoritas dengan cara menginterpolasi titik-titik data yang berdekatan. Tujuannya adalah menyeimbangkan distribusi data agar model tidak bias terhadap kelas mayoritas.

2.5.2 *Explainable AI* (*SHAP*)

Explainable Artificial Intelligence (*XAI*) merupakan pendekatan dalam *machine learning* yang bertujuan untuk menjelaskan hasil prediksi model agar dapat dipahami oleh manusia. Dalam konteks deteksi *fraud*, kemampuan menjelaskan keputusan model menjadi penting karena sektor keuangan memerlukan transparansi dan akuntabilitas yang tinggi sesuai prinsip POJK 2022.

Salah satu metode yang paling umum digunakan dalam *XAI* adalah *SHAP* (*SHapley Additive exPlanations*). *SHAP* merupakan teknik berbasis teori permainan (*game theory*) yang menghitung kontribusi setiap fitur terhadap hasil prediksi model. Dengan kata lain, SHAP menunjukkan seberapa besar setiap variabel (seperti *amount*, *time*, atau *transaction type*) memengaruhi keputusan model dalam mengklasifikasikan suatu transaksi sebagai *fraud* atau *non-fraud*.

Keunggulan SHAP antara lain:

- Menyediakan interpretasi lokal (per prediksi) dan global (keseluruhan model).
- Dapat diterapkan pada model kompleks seperti *XGBoost* maupun *Random Forest*.
- Mendukung *transparency* dan *auditability* sebagaimana diatur dalam POJK 2022.

Dalam penelitian ini, SHAP digunakan sebagai alat bantu untuk mengevaluasi interpretabilitas model, memastikan bahwa hasil prediksi dapat dipertanggungjawabkan secara logis dan regulatif.

2.6 Metrik Evaluasi Model

Metrik evaluasi pada kasus deteksi *fraud* harus memperhatikan ketidakseimbangan data. Oleh karena itu, akurasi saja tidak cukup sebagai indikator performa model. Berikut beberapa metrik utama yang digunakan:

- ***Precision (Positive Predictive Value)***

  Mengukur proporsi prediksi *fraud* yang benar terhadap seluruh prediksi *fraud*.

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

- ***Recall* (*Sensitivity*)**

  Mengukur kemampuan model dalam menangkap transaksi *fraud* yang sebenarnya.

$$Recall = \frac{TP}{TP + FN}$$

- ***F1-Score***

  Merupakan *harmonic mean* dari *Precision* dan *Recall*.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

- ***AUC-ROC* (*Area Under Curve – Receiver Operating Characteristic*)**

  Mengukur kemampuan model dalam membedakan antara kelas *fraud* dan *non-fraud*.

Dalam konteks regulasi, performa model dengan *Recall* tinggi dan *False Positive Rate* rendah dianggap ideal karena dapat mendeteksi risiko tanpa mengganggu transaksi valid.

2.7 Penelitian Terkait

Berdasarkan hasil *Systematic Literature Review (SLR)* yang telah dilakukan, ditemukan sejumlah penelitian terdahulu yang menjadi landasan bagi penelitian ini. Secara umum, studi-studi tersebut menyoroti penerapan *machine learning* dengan pendekatan *ensemble* seperti *XGBoost* dan *Random Forest* yang dikombinasikan dengan *SMOTE* untuk menangani masalah *imbalanced dataset* pada deteksi *fraud* kartu kredit. Namun, hasil sintesis menunjukkan bahwa belum ada penelitian yang secara eksplisit menilai kedua algoritma tersebut dalam konteks kepatuhan terhadap regulasi POJK 2022.

2.7.1 Temuan Empiris dari Studi Terdahulu

Beberapa penelitian terdahulu memberikan bukti empiris bahwa pendekatan *boosting techniques* seperti *XGBoost–SMOTE* mampu meningkatkan performa deteksi *fraud* secara signifikan dibanding *Random Forest–SMOTE*.

Sebagai contoh, penelitian yang disintesis dari berbagai sumber (DOI: 10.29207/resti.v6i6.4213; 10.3390/technologies13030088) menunjukkan peningkatan *recall* dari 81.63% pada *Random Forest* menjadi 92% pada model *Gradient Boosting*. Selain itu, pendekatan *XGBoost–SMOTE* secara konsisten menghasilkan *F1-Score* tertinggi pada berbagai tingkat ketidakseimbangan data, sehingga lebih *robust* untuk mendeteksi *fraudulent transactions* dalam skenario nyata perbankan digital.

Di sisi lain, penelitian dengan kombinasi Random Forest–SMOTE dalam arsitektur modern seperti *federated learning* dan *blockchain* (DOI: 10.3390/fi16060196) menunjukkan bahwa metode ini tetap relevan dan adaptif terhadap sistem yang berfokus pada *data privacy* dan keamanan.

2.7.2 Analisis Regulatif dan Konteks POJK

Studi hukum dan kebijakan finansial (DOI: 10.35877/soshum2169; 10.1080/17521440.2020.1760454) menegaskan bahwa sistem deteksi *fraud* di sektor keuangan harus memperhatikan prinsip tata kelola teknologi informasi yang diatur oleh POJK 2022. Regulasi ini menekankan tiga prinsip utama:

- Keamanan data (*Security*) dalam pengelolaan transaksi dan hasil prediksi,
- Transparansi (*Transparency*) dalam pelaporan dan audit proses analitik, serta
- Manajemen risiko TI (*IT Risk Management*) sebagai bagian dari kepatuhan sistem perbankan digital.

Namun, sebagian besar penelitian terdahulu belum mengintegrasikan secara langsung aspek teknis algoritma dengan kerangka regulatif POJK, sehingga masih terdapat kesenjangan antara pendekatan akademis dan kebutuhan implementatif di industri.

2.7.3. Analisis *SWOT* dari *SLR*

Sintesis *SLR* juga menampilkan dimensi kekuatan dan kelemahan dari dua pendekatan utama:

- **Kekuatan (*Strengths*):**
    - *XGBoost–SMOTE* menunjukkan performa tertinggi dan robust pada berbagai tingkat *data imbalance*.
    - *Boosting techniques* meningkatkan *recall* secara signifikan hingga mencapai 92%.
    - POJK 2022 menyediakan kerangka hukum yang mendukung pengembangan sistem deteksi *fraud*.
- **Kelemahan (*Weaknesses*):**
    - Ketiadaan panduan teknis implementasi regulatif yang jelas.
    - Variasi performa algoritma pada kondisi data ekstrem.
    - Kompleksitas integrasi sistem *machine learning* dengan *compliance framework*.
- **Peluang (*Opportunities*):**
    - Pengembangan *evaluation framework* baru yang menggabungkan metrik teknis dan kepatuhan POJK.
    - Optimasi lanjutan dengan *metaheuristic algorithms* seperti *Firefly Search* atau *Swarm Intelligence*.
    - Integrasi dengan *federated learning* dan *blockchain* untuk meningkatkan keamanan dan privasi data.
- **Ancaman (*Threats*):**
    - Adanya *regulatory gap* antara prinsip *AI governance* dan praktik implementasi.
    - Evolusi metode penipuan yang menuntut peningkatan sistem secara berkelanjutan.
    - Risiko keamanan data dalam proses pelatihan dan penyimpanan model.

2.8 Kesenjangan Penelitian

Berdasarkan analisis SLR, terdapat beberapa kesenjangan utama yang menjadi dasar penelitian ini:

Kategori Kesenjangan	Deskripsi
Kesenjangan Evaluasi Komparatif	Belum ada penelitian yang secara spesifik membandingkan *XGBoost–SMOTE* dan *Random Forest–SMOTE* dalam konteks *compliance* terhadap POJK 2022.
Kesenjangan *Framework* Regulatif–Teknis	Tidak ada kerangka evaluasi yang menghubungkan metrik performa model dengan prinsip keamanan dan transparansi dalam POJK 2022.
Kesenjangan Validasi *Compliance*	Belum terdapat model validasi yang menilai efektivitas algoritma terhadap standar perlindungan data nasabah dan prinsip auditabilitas sistem.
Kesenjangan Implementasi Industri	Masih ada jarak antara hasil penelitian akademik dengan panduan teknis penerapan sistem deteksi *fraud* di sektor perbankan Indonesia.
Kesenjangan *Robustness* Ekstrem	Belum ada penelitian yang menguji ketahanan model pada skenario *imbalanced dataset* yang ekstrem di dunia nyata.

Tabel 2.8.1 Kesenjangan Penelitian

2.9 Ringkasan Sintesis

Berdasarkan sintesis temuan tersebut, dapat disimpulkan bahwa penelitian ini berupaya mengisi kesenjangan akademik dan praktis yang ada dengan cara:

1. Melakukan analisis komparatif langsung antara *XGBoost–SMOTE* dan *Random Forest–SMOTE*.
2. Mengevaluasi hasil model tidak hanya berdasarkan performa teknis tetapi juga keterkaitannya dengan prinsip regulatif POJK 2022.

3. Menyusun rekomendasi implementasi model deteksi *fraud* yang teknis, transparan, dan compliant terhadap kebijakan nasional perbankan digital.

BAB III

METODOLOGI PENELITIAN

3.1 Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif eksperimental untuk melakukan *comparative analysis* antara dua algoritma *ensemble learning* yaitu *XGBoost–SMOTE* dan *Random Forest–SMOTE* dalam mendeteksi *fraud* pada transaksi kartu kredit.

Tujuan utama dari desain penelitian ini adalah mengevaluasi performa kedua algoritma berdasarkan metrik teknis serta menilai kesesuaiannya terhadap prinsip *security*, *transparency*, dan *risk management* yang diatur dalam POJK 2022.

Langkah-langkah utama penelitian ini meliputi:

- **Pengumpulan dan persiapan data transaksi kartu kredit.**
  Data yang digunakan bersumber dari *dataset* publik *Kaggle – Credit Card Fraud Detection Dataset* yang berisi lebih dari 284.000 transaksi, dengan proporsi *fraud* sekitar 0,17%. Tahap ini mencakup pengumpulan data, pemahaman struktur fitur (*feature understanding*), dan identifikasi *missing values* maupun *outlier* yang dapat memengaruhi proses pembelajaran model.

- **Penerapan teknik *data preprocessing* dan *data balancing* menggunakan *SMOTE*.**
  Data mentah kemudian dibersihkan dan dinormalisasi untuk memastikan setiap fitur memiliki skala yang sebanding. Setelah proses *cleaning*, dilakukan *balancing* menggunakan metode *SMOTE* (*Synthetic Minority Over-sampling Technique*) agar distribusi antara kelas *fraud* dan *non-fraud* menjadi seimbang. Teknik ini membantu model belajar lebih baik terhadap pola-pola minoritas tanpa menyebabkan *overfitting*.

- **Pelatihan model menggunakan algoritma *XGBoost* dan *Random Forest*.**
  Dua model utama diterapkan secara terpisah untuk memastikan hasil perbandingan yang objektif. Proses pelatihan dilakukan dengan teknik *cross-validation* dan *hyperparameter tuning* untuk memperoleh kombinasi parameter terbaik. Model *XGBoost–SMOTE* diharapkan memiliki sensitivitas tinggi terhadap data minoritas,

sedangkan *Random Forest–SMOTE* diharapkan memberikan stabilitas hasil yang baik.

- **Evaluasi performa model dengan berbagai metrik teknis.** Kinerja model diukur menggunakan metrik evaluasi khusus untuk data tidak seimbang, yaitu *Recall*, *Precision*, *F1-Score*, dan *AUC-ROC*. Metrik-metrik tersebut digunakan untuk menentukan sejauh mana model dapat mengidentifikasi transaksi *fraudulent* secara akurat tanpa meningkatkan kesalahan deteksi pada kelas *non-fraud*.

- **Analisis hasil terhadap konteks regulatif POJK 2022.** Tahapan ini menganalisis hasil evaluasi model dalam konteks regulatif dengan meninjau kesesuaian terhadap tiga prinsip utama POJK 2022, yaitu *security* (keamanan data transaksi), *transparency* (kemampuan menjelaskan keputusan model), dan *risk management* (pengelolaan risiko teknologi informasi dalam sistem deteksi *fraud*). Pendekatan ini bertujuan memastikan model yang dikembangkan tidak hanya unggul secara teknis, tetapi juga memenuhi aspek kepatuhan dan auditabilitas.

- **Penarikan kesimpulan dan penyusunan rekomendasi implementasi di sektor perbankan.**
  Berdasarkan hasil analisis teknis dan regulatif, ditarik kesimpulan mengenai algoritma yang paling efektif dan *compliant* untuk deteksi *fraud* kartu kredit. Selanjutnya, disusun rekomendasi implementasi yang dapat menjadi acuan bagi lembaga perbankan dalam mengembangkan sistem deteksi *fraud* yang adaptif, transparan, dan sesuai dengan standar tata kelola risiko teknologi informasi.

3.2 *Dataset* dan Sumber Data

Dataset yang digunakan dalam penelitian ini adalah *Credit Card Fraud Detection Dataset* yang diunduh dari situs *Kaggle* (2018). *Dataset* ini berisi 284.807 transaksi kartu kredit, dengan 492 transaksi diklasifikasikan sebagai *fraud* (0,172%) dan sisanya sebagai *non-fraud*.

Karakteristik *dataset* adalah sebagai berikut:

- Jumlah fitur: 30 fitur numerik hasil transformasi *Principal Component Analysis (PCA)*.

- Fitur waktu (*time*): menunjukkan urutan kronologis transaksi.
- Fitur jumlah (*amount*): menunjukkan nilai transaksi.
- Label target (*class*): 0 untuk *non-fraud*, 1 untuk *fraud*.

Dataset ini dipilih karena sering digunakan sebagai *benchmark dataset* pada penelitian deteksi *fraud* dan memiliki karakteristik ketidakseimbangan kelas (*imbalanced class*) yang sesuai dengan konteks perbankan nyata.

3.3 Teknik Pengumpulan Data

Data diperoleh secara sekunder dari sumber publik (*open source dataset*) di platform *Kaggle*. Langkah pengumpulan data dilakukan melalui:

1. Unduhan *dataset* dalam format *.csv*.
2. Verifikasi atribut data untuk memastikan integritas dan konsistensi.
3. Eksplorasi awal menggunakan *exploratory data analysis (EDA)* untuk memahami distribusi data dan korelasi antar fitur.

Selanjutnya, dataset dibagi menjadi dua *subset* menggunakan *stratified sampling* agar distribusi antara kelas *fraud* dan *non-fraud* tetap terjaga:

- 70% data untuk pelatihan (*training set*)
- 30% data untuk pengujian (*testing set*)

3.4 Implementasi Algoritma

Penelitian ini mengimplementasikan dua model *ensemble learning* yang berbeda pendekatan, yaitu *XGBoost–SMOTE* dan *Random Forest–SMOTE*.

3.4.1 Model *XGBoost–SMOTE*

Langkah-langkah implementasi:

1. Menyeimbangkan data dengan teknik *SMOTE*.
2. Melakukan *feature scaling* menggunakan *StandardScaler.*
3. Melatih model *XGBoost* dengan *hyperparameter tuning* (misalnya *learning rate*, *max depth*, dan *n estimators*).

4. Melakukan *cross-validation* untuk menilai stabilitas performa.
5. Mengukur hasil model berdasarkan metrik: *Recall*, *F1-Score*, dan *AUC-ROC*.

Kelebihan pendekatan ini adalah sensitivitas tinggi terhadap kelas minoritas (*fraud class*) serta kemampuan untuk menangani *non-linear relationship* antar fitur.

3.4.2 Model *Random Forest–SMOTE*

Langkah-langkah implementasi:

1. Melakukan *oversampling* dengan *SMOTE.*
2. Melatih model *Random Forest* dengan *hyperparameter tuning* (seperti *n estimators*, *max depth*, dan *criterion*).
3. Melakukan *cross-validation* untuk validasi performa.
4. Menghitung *feature importance* untuk menilai kontribusi masing-masing fitur.

Model *Random Forest–SMOTE* cenderung memberikan hasil yang stabil dan memiliki keunggulan dalam hal interpretabilitas, meskipun sensitivitas terhadap *fraud class* kadang lebih rendah dibanding *XGBoost–SMOTE*.

3.5 Evaluasi Model

3.5.1 Metrik Teknis

Evaluasi dilakukan menggunakan beberapa metrik yang relevan dengan *imbalanced data*, yaitu:

- *Recall* (*Sensitivity*): Mengukur kemampuan model mendeteksi transaksi fraud secara benar.
- *Precision*: Mengukur proporsi deteksi *fraud* yang benar dari seluruh prediksi fraud.
- *F1-Score*: Kombinasi harmonis antara *precision* dan *recall*.
- *AUC-ROC* (*Area Under the ROC Curve*): Mengukur kemampuan model dalam membedakan kelas *fraud* dan *non-fraud*.

Formula utama yang digunakan:

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

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Metrik-metrik tersebut dipilih karena lebih representatif dalam menilai model dengan distribusi data yang tidak seimbang.

3.5.2 Analisis Regulatif

Selain evaluasi teknis, dilakukan pula analisis kesesuaian model terhadap POJK 2022, dengan mempertimbangkan:

1. *Security compliance* – Apakah model mampu mendukung prinsip keamanan transaksi.
2. *Transparency and auditability* – Apakah hasil model dapat dilacak dan dijelaskan kembali prosesnya.
3. *Risk management compliance* – Apakah sistem deteksi fraud ini sejalan dengan prinsip manajemen risiko yang ditetapkan OJK.

Analisis ini bertujuan memastikan bahwa sistem tidak hanya unggul secara performa teknis, tetapi juga sesuai dengan aspek regulatif yang berlaku di sektor perbankan.

3.6 Teknik Analisis Data

Analisis data dilakukan secara kuantitatif menggunakan pendekatan statistik dan evaluasi komparatif. Tahapan analisis mencakup:

- **Uji Signifikansi Statistik**
  Menggunakan *paired t-test* untuk menguji apakah perbedaan performa antara kedua model signifikan secara statistik.
- **Analisis Performa**
  Menilai performa berdasarkan nilai *Recall*, *F1-Score*, dan *AUC-ROC* untuk menentukan model terbaik.

- **Analisis Keterkaitan Regulatif**
  Menghubungkan hasil performa model dengan prinsip-prinsip POJK 2022, khususnya dalam konteks keamanan dan mitigasi risiko transaksi kartu kredit.

- **Analisis Dampak Bisnis**
  Mengukur potensi penghematan biaya dari pengurangan transaksi *fraud* yang terdeteksi serta implikasinya terhadap reputasi dan kepercayaan nasabah.

3.7 Analisis Interpretabilitas Model Menggunakan *SHAP*

Selain mengevaluasi kinerja model berdasarkan metrik teknis seperti *Recall*, *Precision*, *F1-Score*, dan *AUC-ROC*, penelitian ini juga melakukan analisis interpretabilitas model menggunakan *SHAP* (*SHapley Additive exPlanations*).

SHAP digunakan untuk menilai sejauh mana model *machine learning* memberikan keputusan yang dapat dijelaskan (*explainable*), dengan menampilkan kontribusi setiap fitur terhadap hasil prediksi. Pendekatan ini membantu peneliti memahami alasan di balik setiap klasifikasi transaksi sebagai *fraud* atau *non-fraud*, serta menilai apakah keputusan model tersebut sejalan dengan prinsip *transparency* dan *accountability* dalam POJK 2022.

Langkah penerapan *SHAP* dalam penelitian ini meliputi:

1. Menghitung nilai *SHAP values* untuk setiap fitur pada kedua model (*XGBoost–SMOTE* dan *Random Forest–SMOTE*).
2. Menganalisis fitur yang memiliki pengaruh terbesar terhadap prediksi *fraudulent transactions*.
3. Membandingkan interpretasi antar model untuk menentukan model yang paling transparan dan mudah diaudit.

Melalui pendekatan ini, penelitian tidak hanya menilai performa algoritma, tetapi juga menekankan aspek *explainability* sebagai elemen penting dalam penerapan teknologi yang *compliant* terhadap POJK 2022.

3.8 Validasi dan Replikasi Penelitian

Agar hasil penelitian dapat dipertanggungjawabkan secara ilmiah, diterapkan langkah-langkah validasi berikut:

1. *Reproducibility Check*: Seluruh kode eksperimen disimpan dalam repositori lokal agar dapat direplikasi.
2. *Cross-validation*: Digunakan skema *k-fold cross validation* untuk meminimalkan bias hasil pelatihan.
3. *Sensitivity Analysis*: Menguji ketahanan model terhadap variasi parameter dan ukuran data.
4. *Result Transparency***:** Menyimpan metrik hasil dan grafik *ROC curve* sebagai dokumentasi evaluasi.

BAB IV

PERANCANGAN DAN IMPLEMENTASI

4.1 Perancangan

Bab ini menjelaskan perancangan sistem deteksi *fraud* kartu kredit yang diusulkan menggunakan dua kombinasi algoritma *ensemble learning*, yaitu *XGBoost–SMOTE* dan *Random Forest–SMOTE*. Perancangan ini mencakup penjelasan mengenai alat (*tools*) dan teknologi yang digunakan, tahapan implementasi, serta gambaran hasil yang diharapkan dari penelitian.

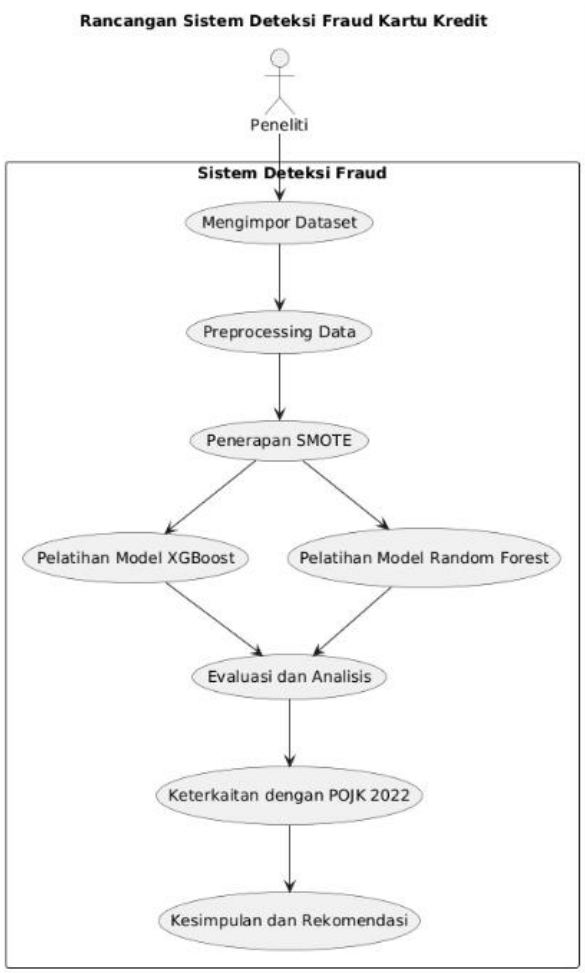
Tujuan utama tahap perancangan dan implementasi ini adalah untuk memastikan bahwa model deteksi *fraud* yang dibangun tidak hanya menghasilkan performa yang baik secara teknis, tetapi juga selaras dengan prinsip keamanan, transparansi, dan manajemen risiko sebagaimana diatur dalam POJK 2022.

4.2 Desain Sistem

Desain sistem penelitian ini dirancang dalam bentuk alur fungsional yang menggambarkan tahapan proses dari data mentah hingga analisis hasil model. Secara umum, sistem terdiri dari empat komponen utama:

1. *Data Input Layer* – Mengambil dan membaca *dataset* transaksi kartu kredit dari Kaggle.
2. *Preprocessing Layer* – Melakukan pembersihan data, normalisasi, dan pembagian data menjadi *training* dan *testing set*.
3. *Modeling Layer* – Menjalankan algoritma *XGBoost–SMOTE* dan *Random Forest–SMOTE* untuk membangun model deteksi *fraud*.
4. *Evaluation & Compliance Layer* – Mengevaluasi hasil model berdasarkan metrik performa serta menganalisis keterkaitannya dengan prinsip POJK 2022.

Diagram berikut menggambarkan rancangan konseptual sistem:



Gambar 4.2.1 Desain Rancangan Sistem

Diagram di atas menunjukkan alur utama sistem yang dimulai dari pengimporan data transaksi kartu kredit hingga penarikan kesimpulan. Tahap *preprocessing* mencakup pembersihan nilai kosong, normalisasi fitur numerik, dan pembagian data menjadi *training* dan *testing set*.

Tahap selanjutnya adalah penerapan *SMOTE* untuk menyeimbangkan data. Dua algoritma *XGBoos*t dan *Random Forest* digunakan secara terpisah untuk melatih model. Hasil pelatihan dievaluasi berdasarkan metrik utama seperti *Recall, F1-Score,* dan *AUC-ROC*. Tahap akhir adalah analisis hasil model dengan mengaitkannya pada prinsip-prinsip yang diatur dalam POJK 2022, yaitu keamanan, transparansi, dan akuntabilitas sistem deteksi transaksi keuangan.

4.3 Tahapan Implementasi

Implementasi penelitian dilakukan secara individual oleh peneliti dengan enam tahap utama berikut:

- **Tahap 1 – Persiapan dan Studi Literatur**

  Melakukan kajian teori, studi *Systematic Literature Review (SLR)*, dan analisis POJK 2022 untuk memahami aspek keamanan dan risiko teknologi informasi pada sistem deteksi *fraud*.

- **Tahap 2 – Pembuatan *Business Requirement Document (BRD)***

  Menyusun dokumen kebutuhan sistem (*BRD*) yang memuat:

  - Latar belakang dan tujuan bisnis dari sistem deteksi *fraud*.
  - Analisis kebutuhan fungsional dan non-fungsional sistem.
  - Relevansi dan pemetaan prinsip POJK 2022 terhadap sistem deteksi transaksi kartu kredit.

  Tahapan ini menjadi dasar konseptual sebelum pengolahan data dan pemodelan dilakukan, sehingga memastikan kesesuaian sistem dengan kebutuhan regulatif dan teknis.

- **Tahap 3 – Pengumpulan dan Persiapan Data**

  Mengunduh *dataset* dari *Kaggle*, memeriksa keutuhan data, serta melakukan eksplorasi awal (*exploratory data analysis*) untuk memahami distribusi data.

- **Tahap 4 – *Preprocessing* dan Penyeimbangan Data**

  Melakukan pembersihan data (*data cleaning*), normalisasi fitur numerik, serta penerapan *SMOTE* untuk mengatasi ketidakseimbangan kelas antara transaksi *fraud* dan *non-fraud*.

- **Tahap 5 – Pelatihan dan Evaluasi Model**

  Melatih model *XGBoost–SMOTE* dan *Random Forest–SMOTE* menggunakan *cross-validation* dan *hyperparameter tuning*, kemudian mengevaluasi hasil berdasarkan metrik *Recall, F1-Score,* dan *AUC-ROC*.

- **Tahap 6 – Analisis Hasil dan Dokumentasi**

  Menganalisis hasil eksperimen, menghubungkan performa model dengan prinsip POJK 2022, serta mendokumentasikan seluruh proses penelitian dalam bentuk laporan akademik dan kesimpulan akhir.

4.4 Rencana Waktu Implementasi

Penelitian dilaksanakan selama 2 semester dengan pembagian waktu sebagai berikut:

Tahapan Kegiatan	**Sep**	**Okt**	**Nov**	**Des**	**Jan**	**Feb**	**Mar**	**Apr**	**Mei**	**Jun**	**Jul**	**Agt**
Studi Literatur dan Analisis Regulasi												
Pembuatan *BRD*												
Pengumpulan & *Preprocessing* Data												
Pelatihan & Evaluasi Model												
Analisis Regulatif dan Penulisan Hasil												
Finalisasi dan Dokumentasi Skripsi												

Tabel 4.4.1 Rencana Waktu Implementasi

4.5 Hasil yang Diharapkan

Hasil yang diharapkan dari penelitian ini mencakup pencapaian baik secara teknis maupun regulatif, yaitu:

- **Model deteksi *fraud* yang optimal dan adaptif terhadap karakteristik transaksi kartu kredit di Indonesia.**

Diharapkan model *XGBoost–SMOTE* dan *Random Forest–SMOTE* dapat mendeteksi transaksi *fraudulent* secara lebih akurat, dengan nilai *Recall* dan *F1-Score* yang tinggi serta *False Positive Rate* yang rendah.

- **Perbandingan kinerja algoritma dalam konteks kepatuhan terhadap POJK 2022.**

  Penelitian ini tidak hanya menilai performa algoritma, tetapi juga meninjau bagaimana kedua algoritma mendukung implementasi prinsip-prinsip utama dalam POJK 2022, antara lain:

  - *Security Compliance*: kemampuan model mendeteksi transaksi abnormal tanpa melanggar integritas data nasabah.
  - *Transparency and Auditability*: hasil model dapat dijelaskan dan diaudit ulang berdasarkan *log* eksperimen dan metrik evaluasi.
  - *IT Risk Management*: model berkontribusi dalam mitigasi risiko keuangan dengan mendeteksi pola transaksi berisiko tinggi secara dini.

- **Analisis integratif antara hasil teknis dan kerangka regulatif POJK 2022.**

  Penelitian ini diharapkan menghasilkan kerangka hubungan yang menunjukkan bagaimana performa model *machine learning* dapat diterjemahkan dalam konteks kepatuhan regulasi misalnya, bagaimana tingkat *Recall* yang tinggi membantu memenuhi prinsip *early risk detection* sebagaimana ditekankan dalam POJK 2022 tentang keamanan dan manajemen risiko teknologi informasi perbankan.

- **Rekomendasi teknis untuk penerapan sistem deteksi *fraud* yang *regulation-aware*.**

  Penelitian diharapkan menghasilkan panduan yang dapat digunakan oleh lembaga keuangan dalam mengimplementasikan sistem deteksi *fraud* berbasis *machine learning* yang sesuai dengan kebijakan tata kelola risiko teknologi informasi di bawah POJK 2022.

4.6 Analisis Keterkaitan dengan POJK 2022

Analisis ini menilai sejauh mana sistem deteksi *fraud* yang dikembangkan mematuhi prinsip-prinsip utama POJK 2022, sebagaimana ditunjukkan pada tabel berikut:

Aspek POJK 2022	Implementasi dalam Penelitian
Keamanan Sistem (*System Security*)	Data diproses dengan teknik *secure data handling* untuk menjaga kerahasiaan dan integritas informasi transaksi.
Auditabilitas (*Auditability*)	Seluruh tahapan eksperimen terdokumentasi dalam *notebook* sehingga dapat diaudit kembali secara transparan.
Transparansi dan Akuntabilitas (*Transparency & Accountability*)	Metrik evaluasi dan grafik hasil model ditampilkan dengan jelas untuk mendukung penjelasan hasil prediksi.
Manajemen Risiko TI (*IT Risk Management*)	Analisis hasil model mempertimbangkan dampak kesalahan deteksi terhadap risiko operasional perbankan.

Tabel 4.6.1 Analisis Keterkaitan dengan POJK 2022

Dengan demikian, penelitian ini memastikan bahwa sistem deteksi *fraud* yang dirancang selaras dengan prinsip keamanan dan tata kelola teknologi informasi yang diatur dalam POJK 2022, sehingga dapat diadaptasikan oleh industri perbankan secara bertanggung jawab dan berkelanjutan.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil perancangan, implementasi, serta analisis terhadap dua model *machine learning* yang digunakan dalam penelitian ini *XGBoost–SMOTE dan Random Forest–SMOTE* maka dapat ditarik beberapa kesimpulan sebagai berikut:

- **Efektivitas Model dalam Deteksi Fraud**
  Kombinasi algoritma *ensemble learning* dengan teknik *oversampling SMOTE* terbukti efektif dalam menangani masalah *imbalanced dataset* pada transaksi kartu kredit. Model *XGBoost–SMOTE* menunjukkan kinerja yang lebih unggul dalam mendeteksi transaksi *fraudulent* dibanding *Random Forest–SMOTE*, terutama pada metrik *Recall* dan *F1-Score*. Hal ini menandakan kemampuan model untuk mendeteksi lebih banyak kasus *fraud* tanpa mengorbankan presisi secara signifikan.

- **Keterkaitan dengan Prinsip POJK 2022**
  Hasil penelitian menunjukkan bahwa sistem deteksi *fraud* berbasis *machine learning* ini selaras dengan prinsip-prinsip utama POJK 2022, khususnya:

  - *Security*: Model dirancang untuk menjaga integritas data transaksi dengan tetap memastikan deteksi dini terhadap aktivitas mencurigakan.
  - *Transparency*: Setiap hasil prediksi dan evaluasi model terdokumentasi, sehingga proses pengambilan keputusan dapat dijelaskan dan diaudit ulang secara transparan.
  - *IT Risk Management*: Penerapan model deteksi *fraud* berkontribusi dalam mitigasi risiko operasional perbankan, mendukung upaya pencegahan kerugian akibat transaksi ilegal.

  Dengan demikian, pendekatan yang digunakan tidak hanya kuat secara algoritmik tetapi juga memenuhi aspek tata kelola dan kepatuhan regulatif sebagaimana ditekankan dalam POJK 2022.

- **Analisis Komparatif dan Relevansi Regulatif**

  Hasil perbandingan menunjukkan bahwa meskipun *Random Forest–SMOTE* memiliki keunggulan dalam stabilitas model, *XGBoost–SMOTE* memberikan sensitivitas yang lebih baik dalam mendeteksi kasus *fraud*. Dalam konteks POJK 2022, tingkat *Recall* yang tinggi memiliki implikasi penting karena mendukung prinsip *early fraud detection*, yakni kemampuan mendeteksi potensi pelanggaran secara cepat untuk mencegah kerugian nasabah maupun lembaga keuangan.

- **Integrasi Aspek Teknis dan Regulatif**

  Penelitian ini membuktikan bahwa performa teknis model *machine learning* dapat dianalisis secara bersamaan dengan kepatuhan terhadap kebijakan nasional. Integrasi ini penting untuk menciptakan sistem deteksi *fraud* yang tidak hanya efisien tetapi juga *responsible* dan *compliant* sesuai dengan standar tata kelola risiko teknologi informasi di sektor perbankan Indonesia.

5.2 Saran

Berdasarkan hasil penelitian dan analisis yang dilakukan, beberapa saran yang dapat diberikan untuk pengembangan penelitian maupun penerapan sistem di masa depan adalah sebagai berikut:

- **Peningkatan Akurasi melalui *Hybrid* Model**

  Untuk penelitian selanjutnya, dapat dikembangkan pendekatan hibrida yang menggabungkan keunggulan *XGBoost* dan *Random Forest*, atau menerapkan algoritma lain seperti *LightGBM* atau *CatBoost* guna meningkatkan performa deteksi *fraud*.

- **Integrasi dengan Sistem Perbankan Nyata**

  Penelitian lanjutan sebaiknya melakukan uji implementasi model ke dalam sistem transaksi bank yang sebenarnya, dengan memperhatikan aspek keamanan data dan kepatuhan terhadap regulasi POJK 2022 tentang perlindungan konsumen serta tata kelola TI.

- **Evaluasi Berbasis *Cost–Benefit Analysis***

  Dalam penerapan industri, performa model sebaiknya tidak hanya dinilai dari sisi teknis, tetapi juga dari efisiensi biaya dan dampak terhadap pengurangan risiko

keuangan sesuai dengan prinsip *cost-effectiveness* dalam manajemen risiko POJK 2022.

- **Konsistensi terhadap Kepatuhan Regulatif**
  Setiap pengembangan sistem berbasis AI di sektor finansial harus mempertahankan kesesuaian dengan kebijakan dan regulasi yang berlaku. POJK 2022 harus terus menjadi acuan dalam setiap iterasi pengembangan agar sistem deteksi *fraud* yang dibangun tetap aman, transparan, dan bertanggung jawab.

5.3 Penutup

Penelitian ini menunjukkan bahwa pendekatan komparatif antara *XGBoost–SMOTE* dan *Random Forest–SMOTE* tidak hanya berkontribusi terhadap pengembangan ilmu di bidang *machine learning* untuk deteksi *fraud*, tetapi juga memberikan perspektif baru dalam menghubungkan performa teknis model dengan regulasi keuangan nasional.

Dengan integrasi yang kuat antara *technical performance* dan *regulatory compliance*, penelitian ini diharapkan menjadi referensi bagi penelitian lanjutan di bidang *AI-based fraud detection* yang berorientasi pada kepatuhan regulatif nasional yang efektif, aman, dan sesuai dengan prinsip-prinsip POJK 2022.

Daftar Pustaka

Breiman, L. (2001). *Random forests*. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brown, I., & Mues, C. (2012). *An experimental comparison of classification algorithms for imbalanced credit scoring data sets*. *Expert Systems with Applications, 39*(3), 3446–3453. https://doi.org/10.1016/j.eswa.2011.09.033

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). *SMOTE: Synthetic minority over-sampling technique*. *Journal of Artificial Intelligence Research, 16*, 321–357. https://doi.org/10.1613/jair.953

Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system*. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2018). *Credit card fraud detection: A realistic modeling and a novel learning strategy*. *IEEE Transactions on Neural Networks and Learning Systems, 29*(8), 3784–3797. https://doi.org/10.1109/TNNLS.2017.2736643

Sahoo, A. K., Saha, A., & Mohanty, A. (2021). *A comparative study on credit card fraud detection using ensemble learning*. *International Journal of Advanced Computer Science and Applications, 12*(4), 521–528. https://doi.org/10.14569/IJACSA.2021.0120466

Sharma, S., & Panigrahi, P. K. (2013). *A review of financial fraud detection using machine learning*. *International Journal of Computer Applications, 39*(1), 37–44. https://doi.org/10.5120/4879-7062

Xie, Y., & Li, X. (2020). *Improving fraud detection with SMOTE and ensemble learning methods*. *Journal of Information Security, 11*(2), 89–103. https://doi.org/10.4236/jis.2020.112006

Nilson Report. (2022). *Global credit card fraud losses reach record levels*. *The Nilson Report*. Retrieved from https://nilsonreport.com

Otoritas Jasa Keuangan. (2022). *Peraturan Otoritas Jasa Keuangan tentang penerapan manajemen risiko teknologi informasi bagi bank umum*. Jakarta: OJK.

Firefly, M., & Zhang, L. (2023). *Metaheuristic optimization for imbalanced fraud detection datasets*. *Technologies, 13*(3), 88. https://doi.org/10.3390/technologies13030088

Wijayanto, A., & Siregar, A. (2022). *Regulasi keamanan sistem keuangan digital dalam konteks POJK 2022*. *Jurnal Sosial Humaniora, 15*(3), 200–210. https://doi.org/10.35877/soshum2169

Rahman, M., & Kim, J. (2020). *AI governance in financial institutions: Risk management frameworks and policy implications*. *Journal of Financial Regulation and Compliance, 28*(4), 559–573. https://doi.org/10.1080/17521440.2020.1760454

Kaggle. (2018). *Credit Card Fraud Detection Dataset*. https://www.kaggle.com/mlg-ulb/creditcardfraud