

Pengembangan Sistem Informasi Akuntansi Berbasis Machine Learning untuk Deteksi Dini Penghindaran Pajak pada Perusahaan Terbuka Indonesia

Moch Rifqi Rizal^{1*)}, Iskandar²⁾

¹⁾Direktorat Jenderal Pajak¹, ²⁾Politeknik Keuangan Negara STAN
m.rifqirizal@gmail.com, iskandar@kemenkeu.go.id

*Corresponding Author

Submitted: 9 Juni 2026

Accepted: 24 Juni 2026

Published: 1 Oktober 2026

ABSTRACT

This research focuses on developing a machine learning-based artificial intelligence architecture to serve as an automated detection instrument for potential tax avoidance among Indonesian publicly traded companies. The exploitation of tax regulation loopholes remains a critical challenge that massively degrades national revenue. The study population comprises corporate entities listed on the Indonesia Stock Exchange (IDX), selected via a purposive sampling approach. Information gathering utilized a dual-method strategy. Secondary data, extracted from historical financial databases within the LSEG Refinitiv Workspace, incorporating parameters such as the Effective Tax Rate (ETR), Cash Effective Tax Rate (CETR), and Debt to Equity Ratio (DER), served as the training dataset for the AI algorithm. Concurrently, primary data was obtained through in-depth interviews with four tax professionals and one academic to qualitatively map out supervisory constraints. The primary methodology involved formulating an anomaly detection model embedded within a web-based prototype to enable real-time assessments. The system's robustness underwent rigorous evaluation through a series of technical trials, encompassing reliability, compliance, recovery, and stress testing. Analytical outcomes demonstrate the model's high proficiency in precisely identifying concealed financial anomalies and suspicious patterns. The implementation of artificial intelligence not only elevates detection accuracy but also drastically accelerates conventional monitoring processes, thereby empowering fiscal authorities to adopt data-driven decision-making practices. Ultimately, the integration of machine learning presents a strategic solution to mitigate tax revenue leakage and optimize state income, while also paving the way for future research explorations concerning the adoption of cutting-edge technology in financial sector governance and tax administration.

Keywords: tax avoidance, effective tax rate, debt to equity ratio, financial statements, taxpayer compliance, accounting information systems

PENDAHULUAN

Berkurangnya penerimaan kas negara dalam skala makro sangat dipengaruhi oleh maraknya praktik penghindaran pajak oleh berbagai perusahaan. Tantangan di Indonesia bertambah berat karena banyak entitas bisnis berskala terbuka secara aktif menekan beban pajaknya melalui eksploitasi celah undang-undang dan skema transaksi yang rumit. Data dari *Tax Justice Network* pada tahun 2020 mengungkapkan bahwa negara berpotensi kehilangan pendapatan hingga USD 4,86 miliar per tahun akibat fenomena ini. Kebocoran yang setara dengan 5,7% dari sasaran pajak nasional ini sebagian besar, atau sekitar USD 4,78 miliar, secara spesifik diakibatkan oleh ulah sektor korporasi.

Dalam menghadapi fenomena tersebut, Direktorat Jenderal Pajak (DJP) sejatinya telah mengupayakan pengawasan ketat melalui implementasi sistem *Compliance Risk Management* (CRM) yang dirancang untuk mengidentifikasi wajib pajak berisiko tinggi (Asmarani, 2020). Namun, terdapat gap penelitian dan kendala operasional yang sangat nyata di lapangan.

berdasarkan wawasan dari praktisi perpajakan, sistem ini masih memiliki kelemahan fundamental terkait akurasi data yang tidak selalu mutakhir sehingga wajib pajak tidak aktif tetap masuk prioritas. Lebih lanjut, terdapat ketidakseimbangan beban kerja yang ekstrem, di mana satu orang petugas pajak harus menangani rasio hingga 10.000 wajib pajak. Keterbatasan metode pengawasan manual ini memunculkan disparitas besar antara volume data perpajakan yang masif dengan kapasitas analisis otoritas pajak, sehingga pola-pola penghindaran pajak yang dinamis kerap gagal terdeteksi secara dini.

Sejumlah literatur terdahulu telah mengkaji akar penyebab munculnya skema penghindaran pajak. Berdasarkan studi-studi tersebut, elemen-elemen fundamental perusahaan seperti tingkat keuntungan, struktur utang, skala bisnis, porsi saham institusi, hingga tren penjualan terbukti secara konsisten memengaruhi keputusan entitas dalam menekan beban pajaknya (Swandi & Prasetyo, 2024). Lebih lanjut, Setyaningsih & Dwi Syamsiah (2024) mencatat bahwa penyalahgunaan mekanisme *transfer pricing* kerap menjadi alat andalan korporasi. Padahal, risiko ini sebenarnya dapat ditekan melalui penerapan tata kelola perusahaan yang ideal dengan melibatkan auditor eksternal dan dewan komisaris independen. Beragam postulat klasik ini memang berhasil membangun fondasi teoretis yang kokoh terkait motif di balik manuver perpajakan. Kendati demikian, problem esensial yang masih menjadi hambatan saat ini adalah bagaimana mengidentifikasi kejanggalan-kejanggalan tersebut di tengah lautan data berskala masif.

Literatur terkini membuktikan bahwa adopsi *machine learning* dan *Natural Language Processing* (NLP) sangat efektif dalam mendeteksi serta memprediksi perilaku kepatuhan wajib pajak. Sebagai contoh, penggunaan model *Artificial Neural Network* (ANN) terbukti mampu memprediksi perilaku pembayaran pajak dengan akurasi mencapai lebih dari 92% (Rosid, 2023). Selain itu, penerapan NLP berhasil meningkatkan akurasi model pengawasan secara signifikan dari 81% menjadi 93% (Iqbalsah, 2023). Model berbasis pohon keputusan seperti *Random Forest* juga menunjukkan keunggulan yang nyata dibandingkan metode konvensional dalam penilaian pajak (Sampa, 2023). Keunggulan algoritma tersebut turut divalidasi dalam keberhasilannya mengidentifikasi kasus-kasus pajak yang penting (Zilli dkk., 2024). Pada pengembangan lain, penggabungan XBRL dengan *machine learning* telah diimplementasikan secara efektif untuk membuat model audit pajak berbasis risiko (Wibowo, 2022). Di sisi lain, algoritma *Support Vector Machine* (SVM) sukses digunakan untuk mendeteksi penghindaran pajak oleh entitas hukum dengan tingkat akurasi yang tinggi (Beiggi & Ayneband, 2024). Secara kolektif, deretan literatur ini menegaskan perlunya pergeseran paradigma menuju solusi kecerdasan buatan terintegrasi yang dapat mengotomatisasi proses deteksi penghindaran pajak tersebut.

Merespons dinamika dan tantangan pengawasan tersebut, penelitian ini menawarkan konsep teoretis berbasis adopsi *grand theory* Deteksi Anomali melalui integrasi menyeluruh Kecerdasan Buatan (AI) dan *Machine Learning* (ML). Teori ini menawarkan pergeseran paradigma dari sistem pengawasan yang menitikberatkan pada analisis faktor secara manual dan intuisi petugas, menuju pendekatan preskriptif yang mengandalkan otomasi serta objektivitas data. Melalui pemanfaatan algoritma cerdas, administrasi perpajakan difasilitasi dengan pengenalan pola komputasional yang mampu mengidentifikasi transaksi dan pelaporan keuangan yang tidak wajar secara otomatis.

Orisinalitas dan kebaruan dalam penelitian ini secara tegas terletak pada pendekatan holistik dalam pengembangan model kecerdasan buataannya. Tidak sekadar berhenti pada tahap pengujian algoritma *machine learning* di atas kertas, riset ini secara inovatif mengimplementasikan dan menanamkan model tersebut ke dalam sebuah prototipe aplikasi antarmuka berbasis web (*TaxSense*) yang memungkinkan dilakukannya analisis anomali secara *real-time* menggunakan data historis perusahaan terbuka di Indonesia. Pendekatan komprehensif ini secara fundamental memperkaya literatur dan menguatkan praktik lapangan dengan menyajikan sistem pengawasan andal, efisien, dan tervalidasi yang mampu mengatasi kelemahan sistem pengawasan perpajakan konvensional.

Berpijak pada fenomena dan urgensi adopsi teknologi yang telah dipaparkan, rumusan masalah dalam penelitian ini difokuskan pada tiga isu teknis utama. Pertama, bagaimana langkah sistematis dalam mengembangkan model kecerdasan buatan berbasis *machine learning* yang andal dengan memanfaatkan data latih dari entitas perusahaan terbuka guna mendeteksi indikasi penghindaran pajak secara presisi. Kedua, bagaimana mekanisme yang tepat dalam mengimplementasikan model

komputasional tersebut agar terintegrasi mulus ke dalam sebuah prototipe *website* pendukung analisis perpajakan. Ketiga, bagaimana mengukur dan menguji kualitas fungsional *website* yang dihasilkan melalui serangkaian evaluasi teknis yang komprehensif.

Selaras dengan rumusan masalah yang diajukan, penelitian ini mengusung tiga tujuan spesifik. Tujuan utamanya adalah untuk memodelkan dan mengembangkan kecerdasan buatan berbasis *machine learning* menggunakan spesifikasi data perusahaan terbuka untuk mendeteksi kecurangan perpajakan secara otomatis. Selanjutnya, penelitian ini bertujuan mengimplementasikan model algoritma tersebut ke dalam sebuah purwarupa web (*website prototype*) guna memfasilitasi otoritas pajak dalam pengawasan *real-time*. Terakhir, penelitian ini bertujuan mengevaluasi kualitas sistem secara keseluruhan melalui serangkaian pengujian teknis yang terukur guna memastikan kelayakannya.

Sebagai kulminasi dari penguatan argumen dan arah riset yang dibangun, penelitian ini mengajukan rumusan hipotesis yang menyatakan bahwa integrasi teknologi *machine learning* mampu meningkatkan tingkat akurasi pendeteksian pola penghindaran pajak secara signifikan dibandingkan dengan metode pengawasan konvensional. Pemanfaatan model pendeteksi anomali yang efisien dan presisi ini diharapkan mampu menjadi katalisator bagi otoritas perpajakan untuk mengambil keputusan yang lebih berbasis data serta pada akhirnya mewujudkan sebuah *data-driven organization*.

STUDI LITERATUR

Landasan Teori

Sebagai instrumen vital organisasi, Sistem Informasi Akuntansi (SIA) beroperasi untuk mengelola data keuangan dari tahap pengumpulan hingga pelaporan. Guna memfasilitasi pengambilan keputusan oleh manajer, investor, pengawas regulasi, maupun *stakeholders* lainnya, SIA sengaja dikembangkan untuk menyediakan informasi yang bernilai ekonomis (Romney & Steinbart, 2015). Di bawah payung sistem informasi manajemen, SIA memegang peran khusus dalam mengeksekusi transaksi keuangan demi mendukung tujuan akuntansi serta mekanisme kendali perusahaan (Wilkinson, dkk, 1999). Struktur komprehensif dari sistem ini, seperti yang diuraikan oleh Hall (2016), mengintegrasikan sumber daya manusia, panduan baku, basis data, infrastruktur teknologi (*hardware* dan *software*), serta sistem pengamanan internal. Seluruh komponen tersebut bersinergi dalam menjalankan fungsi inti, yaitu mengamankan penyimpanan data, memprosesnya menjadi informasi siap pakai, serta mengawal aset organisasi melalui sistem pengendalian internal yang akurat.

Setiap pengembangan sistem, baik itu sistem informasi akuntansi pokok maupun aplikasi pengawasan yang kelak mengolah data dari sistem tersebut, memerlukan validasi akhir agar dapat beroperasi secara efektif. Oleh karena itu, *User Acceptance Test* (UAT) diterapkan sebagai fase final dalam siklus pengujian perangkat lunak yang bertujuan mengevaluasi sejauh mana suatu sistem memenuhi persyaratan fungsional dan ekspektasi pengguna akhir (Hambling & van Goethem, 2013). UAT berfungsi sebagai mekanisme verifikasi untuk memastikan implementasi sistem beroperasi optimal dalam lingkungan aktual. Evaluasi UAT pada *website* menggunakan metode wawancara mendalam untuk mengungkap aspek kontekstual dan pengalaman subjektif yang tidak terdeteksi melalui pengukuran numerik. Evaluasi ini didasarkan pada standar ISO 9241-11:2018 yang menekankan tiga komponen utama *usability*: efektivitas, efisiensi, dan kepuasan pengguna (ISO, 2018). Integrasi wawancara dalam UAT memberikan *insight* kualitatif yang holistik sebagai dasar empiris untuk iterasi desain yang lebih responsif.

Selain keandalan teknis sistem yang diuji melalui UAT, efektivitas pengawasan sangat bergantung pada bagaimana data yang dihasilkan dimanfaatkan untuk merumuskan kebijakan strategis, sebagaimana dijelaskan dalam konsep *Data Driven Decision Making* (DDDM). DDDM adalah pendekatan sistematis dan terukur yang bertujuan menghasilkan keputusan yang objektif, akurat, dan dapat dipertanggungjawabkan, sekaligus mengurangi bias dan spekulasi subjektif (Fawcett & Provost, 2013). Pada era modern, DDDM sangat relevan di bidang perpajakan dengan memanfaatkan AI dan *Big Data* untuk menganalisis anomali penghindaran pajak secara lebih akurat (Brynjolfsson & McElheran, 2016). Penggunaan algoritma *machine learning* sangat esensial dalam DDDM untuk menyaring data, menemukan pola, dan melakukan prediksi. Proses ini bersifat

iteratif dan adaptif terhadap perubahan regulasi (Shmueli dkk., 2017). Meskipun mendukung otomatisasi yang efisien, tantangan utama dari implementasi DDDM adalah memastikan kualitas dan ketersediaan data yang memadai.

Dalam menghadapi kendala pengawasan, otoritas pajak sangat membutuhkan analisis kebijakan yang berlandaskan bukti empiris serta data berkualitas tinggi untuk mengidentifikasi skema penghindaran pajak (*tax avoidance*). Berbeda dari penggelapan pajak (*tax evasion*) yang melanggar hukum, penghindaran pajak merupakan strategi yang sah secara legal untuk menekan beban pajak dengan cara mengeksploitasi kelemahan pada regulasi (Alstadsaeter dkk., 2021). Kendati demikian, manuver agresif semacam ini kerap memicu polemik etis karena dianggap menyimpang dari tujuan awal pembentukan peraturan perundang-undangan (Lenz, 2020).

Dorongan mendasar bagi entitas korporasi dalam menerapkan skema tersebut adalah upaya untuk mengoptimalkan keuntungan bagi para pemegang saham. Fenomena ini turut diperparah oleh berbagai faktor, seperti kerumitan aturan perpajakan, penegakan kebijakan fiskal yang tidak ketat, serta keberadaan korporasi multinasional yang memanfaatkan skema *Base Erosion and Profit Shifting* (BEPS) (OECD, 2016). Beberapa metode yang lazim dipraktikkan mencakup penentuan harga transfer (*transfer pricing*), pengalihan laba (*profit shifting*), kapitalisasi tipis (*thin capitalization*), *hybrid mismatch arrangements*, penundaan pembayaran pajak (*tax deferral*), hingga pemanfaatan negara suaka pajak (*tax havens*). Sebagai langkah antisipasi dan deteksi, otoritas dapat memanfaatkan sejumlah metrik keuangan, antara lain *Effective Tax Rate* (ETR), *Cash Effective Tax Rate* (Cash ETR), rasio utang terhadap modal (*Debt to Equity Ratio*/DER), serta rasio Laba Sebelum Pajak terhadap Total Pendapatan (*PBT to Revenue Ratio*) untuk melacak potensi penekanan nilai pelaporan pendapatan.

Mengingat rumitnya skema serta banyaknya indikator rasio keuangan yang harus dianalisis secara bersamaan, proses pengawasan manual seringkali tidak lagi memadai, sehingga pendekatan modern mulai mengadopsi *Artificial Intelligence* (AI) atau Kecerdasan Buatan. AI adalah ilmu komputer yang mengembangkan sistem untuk meniru kecerdasan manusia, termasuk mempelajari pola, memproses bahasa alami, dan membuat keputusan otomatis (Russell dkk., 2020). Mesin dilatih menggunakan algoritma canggih untuk memproses *big data* secara efisien dan memprediksi hasil (Mitchell, 1997). Cabang utama AI yang digunakan dalam ranah ini meliputi *Machine Learning* (baik *supervised* maupun *unsupervised learning*), *Natural Language Processing* (NLP), dan *Deep Learning* yang menggunakan jaringan saraf tiruan berlapis untuk pola yang lebih kompleks (LeCun dkk., 2015). Keunggulan AI dalam pajak adalah kemampuannya menganalisis data secara *real-time* dengan efisiensi tinggi, meskipun sistem ini sangat bergantung pada kualitas data latihnya.

Dalam penerapannya untuk membedakan transaksi wajar dan tidak wajar pada kasus perpajakan, AI secara spesifik beroperasi berlandaskan Teori Deteksi Anomali (*Anomaly Detection Theory*). Deteksi anomali berfokus pada identifikasi kejadian atau perilaku yang berbeda signifikan dari pola normal, yang sangat krusial dalam mendeteksi penipuan keuangan atau penghindaran pajak (Chandola dkk., 2009). Anomali terbagi menjadi tiga: anomali titik (*point*), anomali kontekstual, dan anomali kolektif. Pendekatan *unsupervised learning* seperti algoritma *k-Means* dan *Autoencoders* sering digunakan karena kemampuannya mendeteksi anomali pada data berdimensi tinggi yang belum berlabel (Hodge & Austin, 2004). Metode berbasis jarak (*proximity-based*) seperti *k-NN* juga umum digunakan. Tantangan utama dalam teori ini adalah tingginya tingkat deteksi positif palsu (*false positives*) dan kebutuhan adaptasi model secara berkelanjutan terhadap dinamika dunia bisnis.

Penelitian Terdahulu

Tabel 1. Penelitian Terdahulu

No	Penulis (Tahun)	Judul Penelitian	Hasil Penelitian	Kategori
1	Swandi & Prasetyo (2024)	Meta Analisis Determinan Penghindaran Pajak	Profitabilitas, <i>leverage</i> , ukuran perusahaan, kepemilikan institusional, dan pertumbuhan penjualan memiliki pengaruh signifikan terhadap praktik penghindaran pajak secara nasional.	Praktik Penghindaran Pajak
2	Devi dkk. (2022)	Analisis Penyebab Penghindaran Pajak (<i>Tax Avoidance</i>) Dalam Laporan Keuangan Pada Perusahaan Yang Terdaftar di BEI Tahun 2016-2019	Profitabilitas dan likuiditas berpengaruh signifikan terhadap praktik penghindaran pajak, sementara <i>leverage</i> dan Komite Audit tidak menunjukkan pengaruh yang signifikan pada perusahaan manufaktur di BEI.	Praktik Penghindaran Pajak
3	Shakil & Tasnia (2022)	<i>Artificial intelligence and tax administration in Asia and the Pacific</i>	Penerapan kecerdasan buatan (AI) terbukti dapat meningkatkan efisiensi dan transparansi dalam administrasi pajak, meskipun faktor biaya implementasi teknologinya masih menjadi tantangan utama.	Potensi Penerapan AI
4	Yalamati (2023)	<i>Identify fraud detection in corporate tax using Artificial Intelligence advancements Solutions Architect</i>	AI sangat efektif untuk mendeteksi penipuan pajak korporasi. Model AI yang digunakan terbukti mampu mengidentifikasi pola-pola penipuan yang lebih kompleks dan adaptif terhadap perubahan strategi wajib pajak.	Potensi Penerapan AI
5	Rosid (2023)	<i>Artificial Neural Networks for predicting taxpaying behaviour of Indonesian firms</i>	Penggunaan model <i>Artificial Neural Networks</i> (ANN) <i>Multilayer Perceptron</i> (MLP) menghasilkan akurasi prediksi perilaku kepatuhan pajak korporasi di Indonesia di atas 92%, dengan margin laba operasi dan pendapatan lain sebagai penentu utama.	Penerapan Model <i>Machine Learning</i>
6	Beiggi & Ayneband (2024)	<i>Detecting Tax Evasion of Legal Entities Using Artificial Intelligence</i>	Model kecerdasan buatan (jaringan saraf tiruan, <i>support vector machine</i> , dan algoritma koloni lebah buatan) sukses mendeteksi penghindaran pajak entitas hukum dengan akurasi pengujian mencapai 83,79%.	Penerapan Model <i>Machine Learning</i>

7	Bhattacharya & Imam (2024)	<i>Mandated CSR spending and Tax aggressiveness: A machine learning-driven analysis</i>	Melalui analisis <i>Conditional Inference Tree</i> (CIT), ditemukan bahwa agresivitas pajak korporasi menurun seiring dengan peningkatan pengeluaran CSR yang diwajibkan oleh regulasi negara.	Penerapan Model <i>Machine Learning</i>
8	Battagliini dkk. (2024)	<i>Refining public policies with machine learning: The case of tax auditing</i>	Pengembangan metode <i>Bayesian</i> hibrida berbasis proses difusi <i>Langevin</i> berhasil meningkatkan akurasi dan efisiensi komputasi dalam memperkirakan model variabel laten berdimensi tinggi untuk kebutuhan audit perpajakan.	Penerapan Model <i>Machine Learning</i>
9	Guenther dkk. (2023)	<i>How Useful Are Tax Disclosures in Predicting Effective Tax Rates? A Machine Learning Approach</i>	Penggunaan <i>Explainable AI</i> berbasis nilai Shapley berhasil mengidentifikasi pos-pos pengungkapan laporan keuangan yang paling berguna dan memiliki kontribusi tinggi untuk memprediksi tarif pajak efektif (<i>Effective Tax Rate</i> / ETR) satu tahun ke depan.	Penerapan Model <i>Machine Learning</i>
10	Wibowo (2022)	<i>XBRL Open Information Model for Risk Based Tax Audit using Machine Learning</i>	Pengembangan sistem berbasis model <i>Random Forest</i> dengan standar data XBRL mampu melakukan klasifikasi risiko wajib pajak secara presisi, sehingga mempermudah alokasi sumber daya pemeriksaan pajak secara efisien.	Penerapan Model <i>Machine Learning</i>

Sumber: Diolah Penulis

Berdasarkan tinjauan terhadap berbagai literatur dan penelitian terdahulu, masih terdapat celah penelitian terkait implementasi praktis kecerdasan buatan dalam ranah pengawasan pajak korporasi. Oleh karena itu, posisi kebaruan atau *novelty* dan kontribusi utama dari penelitian ini dapat ditegaskan secara eksplisit melalui empat aspek utama. Pertama, penelitian ini secara inovatif melakukan integrasi model *Autoencoder-based anomaly detection* secara langsung ke dalam kerangka sistem informasi akuntansi perpajakan. Kedua, studi ini tidak sekadar berhenti pada pengujian algoritma secara teoretis, melainkan menghadirkan implementasi pendeteksian anomali secara *real-time* melalui aplikasi purwarupa fungsional yang *web-based*.

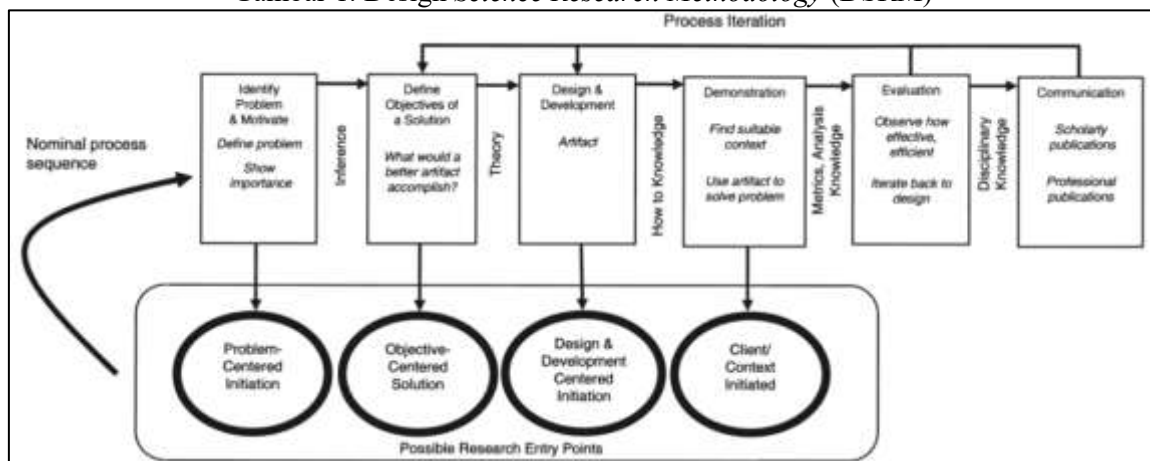
Selanjutnya, aspek kebaruan ketiga dari penelitian ini terletak pada pendekatan evaluasi yang ditawarkan. Studi ini memberikan kombinasi pengujian yang komprehensif dengan memadukan evaluasi metrik kinerja model *machine learning* secara kuantitatif dan evaluasi *usability* sistem berbasis standar ISO 9241-11 secara kualitatif. Keempat, penelitian ini secara spesifik menerapkan pendekatan *unsupervised learning* pada kasus penghindaran pajak oleh entitas perusahaan terbuka di Indonesia. Pendekatan komputasional ini secara strategis digunakan untuk mengatasi kendala keterbatasan data berlabel pada metode pengawasan konvensional, sehingga mampu menghasilkan instrumen deteksi yang lebih adaptif dan objektif.

METODE

Penelitian ini menerapkan kerangka kerja *Design Science Research Methodology* (DSRM) untuk memandu alur kerja komputasi cerdas. Siklus riset ini diawali dengan tahap identifikasi masalah untuk menganalisis kelemahan sistem pengawasan pajak konvensional dan kesenjangan beban kerja otoritas pajak, dilanjutkan dengan merumuskan spesifikasi teknis model deteksi anomali perpajakan berbasis kecerdasan buatan (*Artificial Intelligence* / AI) dan prototipe aplikasi secara *real-time*. Pada tahap berikutnya, dilakukan perancangan dan pengembangan model *Deep*

Clustering menggunakan *Autoencoder*, model *Deep Prediction* berbasis *Keras Sequential*, serta antarmuka web *TaxSense*. Setelah itu, prototipe tersebut didemonstrasikan untuk memproses input laporan keuangan perusahaan secara interaktif, sebelum akhirnya dievaluasi kinerjanya secara kuantitatif melalui metrik statistika serta pengujian kualitas perangkat lunak komprehensif, dan dikomunikasikan dalam bentuk publikasi ilmiah ini.

Gambar 1. Design Science Research Methodology (DSRM)



Sumber: Peffers dkk. (2007)

Populasi sasaran dalam penelitian ini mencakup seluruh entitas perusahaan terbuka yang terdaftar di Bursa Efek Indonesia dengan periode pemutakhiran data historis hingga awal tahun 2025. Pengambilan sampel dilakukan menggunakan metode *purposive sampling* dengan kriteria inklusi yang ketat. Kriteria tersebut meliputi kelengkapan publikasi laporan keuangan tahunan pada platform *London Stock Exchange Group (LSEG) Refinitiv Workspace*, memiliki kelengkapan data historis pasca tahun 1990 untuk menghindari presentasi nilai kosong yang tinggi, serta tidak mengandung nilai nol pada rasio keuangan krusial guna mencegah distorsi komputasional. Melalui kriteria pembersihan data tersebut, diperoleh *dataset* akhir yang representatif sebanyak 10.052 entri data historis laporan keuangan yang siap digunakan dalam fase pemodelan kecerdasan buatan.

Dalam studi ini, variabel keuangan dan perpajakan diukur melalui pemanfaatan rasio keuangan standar yang memiliki korelasi teoretis dengan struktur modal korporasi dan kecenderungan praktik penghindaran pajak. Indikator pengukuran tersebut melibatkan *Effective Tax Rate (ETR)* sebagai representasi proporsi beban pajak berjalan dan tangguhan dibandingkan dengan laba sebelum pajak, diikuti oleh *Cash Effective Tax Rate (CETR)* guna meninjau aliran kas aktual yang disetorkan untuk melunasi kewajiban perpajakan. Di samping itu, riset ini juga mengikutsertakan *Debt to Equity Ratio (DER)* untuk menilai porsi utang terhadap ekuitas, serta *Profit Before Tax (PBT)* yang difungsikan sebagai variabel kontrol untuk memantau skala profitabilitas entitas. Selanjutnya, penentuan status anomali perpajakan dikonversi ke dalam klasifikasi biner. Angka satu melambangkan adanya indikasi kuat penghindaran pajak, sedangkan angka nol menunjukkan aktivitas pelaporan transaksi yang wajar. Pelabelan target biner ini diformulasikan berdasarkan penetapan ambang batas *reconstruction error* yang diproses melalui algoritma *Autoencoder*.

Data yang digunakan dalam riset ini menggabungkan pendekatan metode campuran antara kuantitatif dan kualitatif. Data sekunder kuantitatif berupa metrik keuangan historis bersumber langsung dari basis data *LSEG Refinitiv Workspace* yang diekstraksi secara sistematis menggunakan fungsionalitas *Screeener* dan kemudian disimpan dalam basis data lokal. Di sisi lain, data primer kualitatif dikumpulkan melalui teknik wawancara mendalam semi-terstruktur bersama lima narasumber ahli yang terdiri atas empat praktisi dari Direktorat Jenderal Pajak dan satu akademisi senior di bidang akuntansi perpajakan. Wawancara ini dipadukan dengan pengujian terpandu di mana narasumber diberikan akses langsung ke prototipe aplikasi *TaxSense* untuk melakukan simulasi *input* data, yang tanggapannya kemudian direkam dan dianalisis menggunakan metode analisis konten berdasarkan kriteria kegunaan sistem standar ISO 9241-11.

Proses analisis data kuantitatif diimplementasikan menggunakan arsitektur *Machine Learning* dan *Deep Learning* berbasis pustaka *TensorFlow* di lingkungan *Python*. Sebelum melatih jaringan saraf tiruan, dilakukan pra-pemrosesan data yang diawali dengan pembersihan nilai kosong dan rekaman bernilai nol. Standardisasi fitur dilakukan menggunakan metode penskalaan *Z-Score* agar seluruh rasio keuangan memiliki skala yang seragam dengan rata-rata nol dan standar deviasi satu. Mengingat kasus anomali pajak secara alamiah jauh lebih sedikit dibandingkan kasus normal, ketidakseimbangan kelas diatasi dengan mengaplikasikan teknik *oversampling Synthetic Minority Over-sampling Technique* (SMOTE) pada data latih guna membangkitkan data sintesis baru di sepanjang garis lurus yang menghubungkan sampel minoritas dengan tetangga terdekatnya secara acak.

Pada tahap pemodelan deteksi anomali tanpa pengawasan, algoritma *Autoencoder* digunakan untuk mengompresi dimensi data *input* ke ruang dimensi rendah melalui *encoder* dengan fungsi aktivasi ReLU, sebelum direkonstruksi kembali oleh *decoder* ke dimensi aslinya menggunakan fungsi aktivasi linear. Model ini dilatih untuk meminimalkan tingkat kesalahan kuadrat rata-rata rekonstruksi, di mana ambang batas anomali ditetapkan secara adaptif pada persentil ke-90 dari distribusi kesalahan tersebut. Setelah label anomali terbentuk secara objektif, klasifikasi terarah dibangun menggunakan model *supervised deep learning* dengan optimasi *hyperparameter* otomatis menggunakan *Keras Tuner* melalui algoritma *RandomSearch* untuk menentukan jumlah unit, fungsi aktivasi, tingkat dropout, dan laju pembelajaran terbaik. Kinerja keseluruhan model ini kemudian dievaluasi menggunakan metrik statistika yang komprehensif, seperti Davies-Bouldin Index, Calinski-Harabasz Index, Silhouette Score, tingkat presisi, sensitivitas, F1-score, serta tingkat diskriminasi model melalui nilai area di bawah kurva karakteristik operasi penerima (ROC-AUC).

HASIL

Hasil penelitian ini disajikan secara sistematis dengan mengintegrasikan temuan kuantitatif dari pengembangan model *machine learning* dan temuan kualitatif dari evaluasi pengguna. Solusi yang dihasilkan adalah Model *Deep Clustering* berbasis *Autoencoder* dan prototipe aplikasi web *TaxSense* yang telah melalui serangkaian validasi teknis dan pengujian penerimaan pengguna.

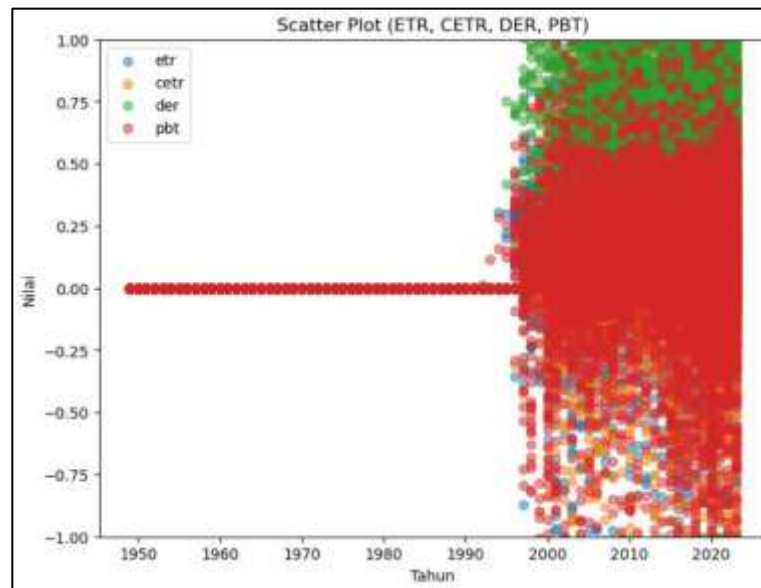
Pemahaman dan Persiapan Data

Tahap awal riset, Pemahaman dan Persiapan Data, berfokus pada laporan keuangan historis perusahaan terbuka BEI yang diekstraksi dari LSEG *Refinitiv Workspace*. Pembersihan data merupakan langkah krusial untuk memastikan *Quality Assurance* data. Hal ini melibatkan penghilangan data sebelum tahun 1990 dan catatan dengan nilai nol pada rasio keuangan (ETR, CETR, DER, dan PBT) guna mengatasi distorsi komputasional dan memperoleh *dataset* yang representatif.

Dalam praktiknya, data mentah (*raw data*) historis sering kali masih mengandung berbagai inkonsistensi yang tidak dapat langsung diolah oleh algoritma *machine learning*. Sebelum melatih jaringan saraf tiruan, pra-pemrosesan mutlak diperlukan karena keberadaan data yang tidak valid, kosong, atau bernilai nol pada metrik keuangan vital akan memicu distorsi komputasional yang signifikan. Jika tidak ditangani secara sistematis, hal ini berisiko menyesatkan model kecerdasan buatan saat proses pengenalan pola, sehingga dapat menurunkan tingkat presisi dalam mendeteksi anomali pada tahap selanjutnya.

Untuk mengidentifikasi seberapa besar skala distorsi tersebut dan menentukan batasan pemotongan data (*data trimming*) yang objektif, langkah analitis pertama yang dilakukan adalah memvisualisasikan seluruh populasi data mentah ke dalam grafik pencar (*scatter plot*). Melalui pemetaan rasio-rasio keuangan terhadap dimensi waktu, anomali struktural pada rentang periode tertentu dapat dievaluasi dengan lebih jelas. Urgensi dari tahapan pembersihan ini dapat dipahami secara komprehensif dengan melihat kondisi persebaran nilai data sebelum diproses, sebagaimana ditunjukkan pada visualisasi berikut:

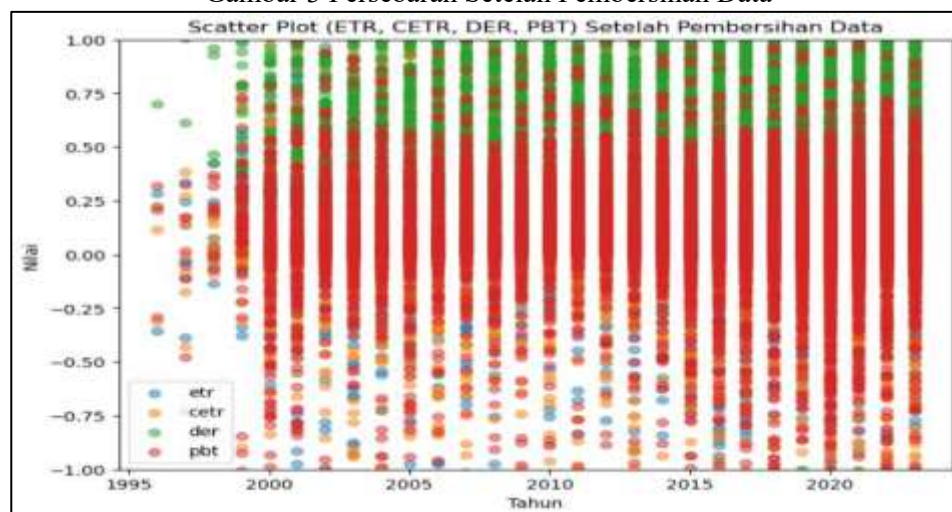
Gambar 2 Persebaran Nilai Data Awal



Sumber: Diolah Penulis

Gambar 2 Persebaran Nilai Data Awal menunjukkan visualisasi awal hubungan antara variabel ETR, CETR, DER, dan PBT terhadap waktu. Grafik ini secara jelas mengidentifikasi bahwa data sebelum tahun 1990 memiliki nilai *null* atau tidak terdefinisi, sehingga tidak relevan untuk analisis inferensial. Oleh karena itu, data pada rentang waktu tersebut harus dihapus, sementara analisis difokuskan pada periode setelah tahun 1990, di mana distribusi data mulai terlihat lebih bervariasi.

Gambar 3 Persebaran Setelah Pembersihan Data



Sumber: Diolah Penulis

Proses pembersihan data yang menghilangkan data sebelum 1990 serta data bernilai nol kemudian menghasilkan sebaran yang lebih representatif, sebagaimana ditunjukkan dalam Gambar 3 Persebaran Setelah Pembersihan Data. Visualisasi ini menegaskan bahwa *dataset* akhir kini tidak lagi terdistorsi oleh nilai ekstrem yang tidak bermakna, memungkinkan interpretasi yang lebih akurat terhadap tren dan hubungan antar rasio keuangan sepanjang periode pengamatan.

Pemodelan dan Hasil Kuantitatif

Dalam tahap Pemodelan, *dataset* dibagi dengan proporsi 90% untuk pelatihan dan 10% untuk pengujian. Model *Autoencoder* dipilih sebagai pendekatan *Deep Learning* untuk deteksi anomali. Setelah penyesuaian *hyperparameter* melalui *Keras Tuner* dan penerapan *EarlyStopping* untuk

mencegah *overfitting*, performa model diukur menggunakan *Reconstruction Error* dan dievaluasi dengan metrik yang komprehensif. Temuan utama penelitian ini terlihat dengan jelas pada metrik evaluasi model *klustering*. Hasil tersebut mendukung hipotesis bahwa integrasi *machine learning* mampu meningkatkan akurasi deteksi secara signifikan.

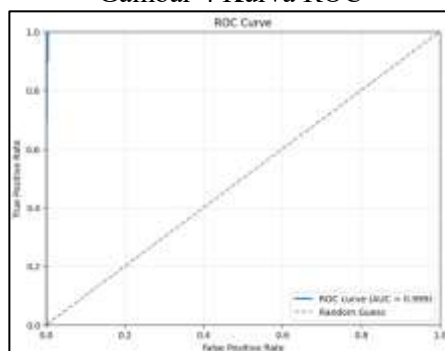
Tabel 2 Skor Metrik Evaluasi Model *Clustering*

No	Evaluasi	Skor
1	Davies-Bouldin Index (DBI)	4.43
2	Calinski-Harabasz Index	214.03
3	Silhouette Score	0.94
4	F1 Score	0.84
5	Precision Score	0.88
6	AUC Score	0.99

Sumber: Diolah Penulis

Tabel 2 Skor Metrik Evaluasi Model *Clustering* menunjukkan kinerja yang baik dari model dengan nilai *ROC-AUC Score* sebesar 0,99, yang mengindikasikan kemampuan diskriminasi yang tinggi dalam membedakan antara data normal dan anomali. Nilai *Silhouette Score* 0,94 menegaskan bahwa struktur kluster yang terbentuk sangat kohesif, menandakan setiap data memiliki kedekatan tinggi dengan pusat kluster masing-masing dan terpisah jauh dari kluster lain. Lebih lanjut, *Precision Score* 0,88 memastikan rendahnya risiko *false positive*, yang berarti sebagian besar data yang diklasifikasikan sebagai anomali memang benar merupakan anomali, serta nilai *F1 Score* 0,84 menunjukkan keseimbangan yang baik antara *precision* dan *recall* dalam konteks data yang tidak seimbang. Sementara itu, metrik *Davies-Bouldin Index* (DBI) sebesar 4,43 dan *Calinski-Harabasz Index* (CHI) sebesar 214,03 menunjukkan pemisahan kluster yang cukup signifikan dan homogenitas yang tinggi dalam setiap kluster, mendukung validitas model dalam mendeteksi pola anomali.

Gambar 4 Kurva ROC



Sumber: Diolah Penulis

Visualisasi Gambar 4 Kurva ROC memperkuat hasil kuantitatif tersebut dengan menampilkan hubungan antara *False Positive Rate* dan *True Positive Rate* pada berbagai titik ambang. Grafik ini memvisualisasikan nilai ROC-AUC yang sangat tinggi (0,999) dan menunjukkan bahwa model memiliki sensitivitas yang andal di setiap level pengambilan keputusan, memastikan kemampuan model dalam memberikan kinerja klasifikasi yang sangat andal untuk deteksi indikasi penghindaran pajak.

Deployment dan Antarmuka Sistem *TaxSense*

Pada tahap *Deployment*, model terbaik diintegrasikan ke dalam sistem prototipe berbasis web untuk menghasilkan prediksi *real-time*. Ambang batas (*threshold*) deteksi anomali ditetapkan pada persentil ke-90 dari distribusi *reconstruction error* data pelatihan, di mana data yang nilai kesalahannya melebihi ambang batas ini otomatis dilabeli sebagai anomali (*ground truth*), memfasilitasi *data-driven decision making*.

Gambar 5 Formulir Isian Website



Sumber: Diolah Penulis

Integrasi ini menghasilkan antarmuka pengguna *TaxSense*, dimulai dengan Gambar 5 Formulir Isian Website. Formulir ini dirancang untuk menerima *input* data keuangan perusahaan secara komprehensif, mencakup variabel seperti Nama Perusahaan, Sektor, Tahun, Total Penghasilan Kotor, Penghasilan Sebelum Pajak, Beban Pajak, Pajak Dibayar, dan Total Aset. Desain yang berpusat pada pengguna ini memastikan proses *penginputan* data menjadi intuitif dan relevan bagi kebutuhan analisis auditor pajak.

Gambar 6 Tampilan Skor Indikasi Penghindaran Pajak



Sumber: Diolah Penulis

Setelah data *diinput* dan diproses, hasil analisis ditampilkan melalui Gambar 6 Tampilan Skor Indikasi Penghindaran Pajak. Tampilan ini menyajikan *output* dalam bentuk Skor Indikasi Penghindaran Pajak (misalnya, 51.8090903759) secara cepat dan terukur, lengkap dengan ID dan versi model yang digunakan. Penyajian skor risiko yang terukur ini memungkinkan otoritas pajak mengidentifikasi dan memprioritaskan entitas yang berisiko tinggi secara instan.

Evaluasi Kualitas Teknis Prototipe

Tahap akhir adalah Evaluasi Kualitas Teknis komprehensif, di mana berbagai pengujian non-fungsional, termasuk *reliability*, *compliance*, *recovery*, dan *stress testing*, berhasil divalidasi, menunjukkan sistem memiliki ketahanan tinggi.

Gambar 7 Rangkuman Data *Reability Testing*

Label	# Samples	Average	Min	Max	Std. Dev.	Error %	Throughput	Received K...	Sent KB/sec	Avg. Bytes
HTTP Requ...	50	150	131	282	22.25	0.00%	50.9/min	47.58	0.20	57432.6
TOTAL	50	150	131	282	22.25	0.00%	50.9/min	47.58	0.20	57432.6

Sumber: Diolah Penulis

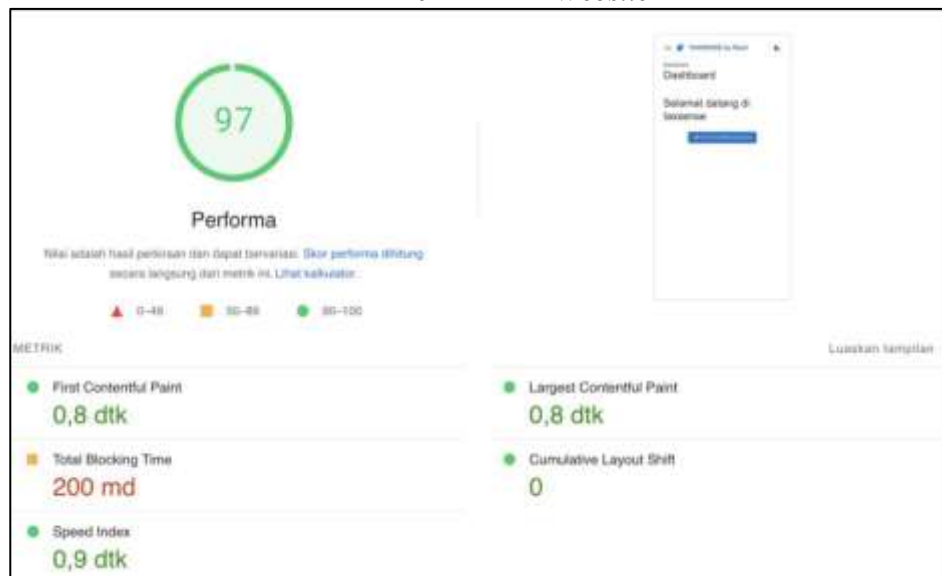
Gambar 7 Rangkuman Data *Reability Testing* menunjukkan bahwa sistem sangat stabil dalam menangani beban kerja normal. Dengan 50 sampel permintaan, sistem mencatatkan waktu respons rata-rata 150 ms dan tingkat kesalahan 0.00%. Kestabilan ini ditegaskan oleh standar deviasi yang sangat kecil (22.25 ms), menunjukkan sistem andal dan konsisten dalam memproses permintaan.

Gambar 8 Rangkuman Sampel *Stress Testing*

Label	# Samples	Average	Min	Max	Std. Dev.	Error %	Throughput	Received...	Sent KB/sec	Avg. Bytes
HTTP Req...	1000	36803	2950	433969	34778.82	18.20%	2.3/sec	106.14	0.22	47200.8
TOTAL	1000	36803	2950	433969	34778.82	18.20%	2.3/sec	106.14	0.22	47200.8

Sumber: Diolah Penulis

Gambar 7 Rangkuman Sampel *Stress Testing* menguji kinerja sistem dalam kondisi ekstrem dengan mengirimkan 1000 HTTP *request* dalam beban tinggi. Meskipun menunjukkan degradasi performa (waktu respons rata-rata 36.803 ms) dan tingkat kesalahan 18,20%, sistem tetap lolos pengujian karena mampu beroperasi tanpa mengalami kegagalan total, membuktikan ketahanan sistem dalam menangani lonjakan beban yang signifikan.

Gambar 9 Performa *Website*

Sumber: Diolah Penulis

Gambar 9 Performa *Website* dari *Volume Testing*, yang diukur menggunakan *Google PageSpeed Insights*, menunjukkan skor performa unggul sebesar 97. Metrik ini didukung oleh kecepatan muat yang impresif, seperti *First Contentful Paint* 0,8 detik dan *Speed Index* 0,9 detik, yang secara keseluruhan menegaskan kapabilitas aplikasi dalam menangani volume akses pengguna yang tinggi dengan responsivitas optimal.

Validasi Kualitatif dan *User Acceptance Test* (UAT)

Selain hasil kuantitatif, wawancara mendalam dengan empat praktisi perpajakan dan satu akademisi menyediakan data primer kualitatif yang krusial. Secara umum, para ahli mengungkapkan bahwa tantangan terbesar dalam pengawasan adalah keterbatasan metode manual dalam menghadapi volume data yang masif dari perusahaan terbuka, diperburuk oleh strategi penghindaran pajak yang semakin kompleks. Seluruh ahli menyepakati bahwa sistem berbasis AI,

seperti *TaxSense*, memiliki potensi transformatif dalam memberikan skor risiko yang objektif dan konsisten.

Hasil dari *User Acceptance Test* (UAT), yang dilakukan sesuai standar ISO 9241-11, mengonfirmasi penerimaan sistem oleh pengguna akhir (praktisi dan akademisi). Penerimaan ini dievaluasi melalui tiga aspek utama yang dirangkum dalam tabel berikut:

Tabel 3 Hasil *User Acceptance Test* (UAT)

Aspek Penilaian	Deskripsi Hasil
Efektivitas	Sistem dinilai sangat efektif dalam mencapai tujuan utamanya, yaitu mengidentifikasi pola anomali secara akurat, serta mampu mengekstraksi wawasan dari data keuangan yang rumit menjadi indikasi risiko yang jelas.
Efisiensi	Terdapat peningkatan efisiensi yang signifikan. Proses analisis manual yang memakan waktu lama kini dapat diselesaikan secara <i>real-time</i> , memungkinkan otoritas pajak merespons temuan anomali dengan lebih cepat.
Kepuasan Pengguna	Desain <i>TaxSense</i> yang intuitif dan visualisasi yang menyederhanakan interpretasi data keuangan menghasilkan kepuasan yang kuat. Sistem juga mampu memberikan wawasan yang dapat segera ditindaklanjuti.

Sumber: Diolah Penulis

Temuan kualitatif ini menegaskan bahwa prototipe *TaxSense* tidak hanya unggul secara metrik teknis (ROC-AUC 0,99), tetapi juga secara fungsional dan *usability*, menjadikannya solusi yang responsif terhadap kebutuhan praktis otoritas perpajakan.

PEMBAHASAN

Pembahasan ini berfungsi sebagai interpretasi mendalam terhadap capaian sistem *TaxSense*, mengaitkan performa yang kompetitif dari model *Deep Clustering* berbasis *Autoencoder* dengan landasan teoretis, serta membandingkannya secara kritis dengan temuan riset terdahulu. Keberhasilan penelitian ini terletak pada integrasi validasi kuantitatif dan kualitatif, yang secara kolektif memvalidasi hipotesis dan menegaskan urgensi penerapan Teori Deteksi Anomali sebagai fondasi bagi administrasi perpajakan berbasis *Data Driven Decision Making* (DDDM).

Interpretasi Kinerja Model *Autoencoder* dan Keunggulan *Unsupervised Learning*

Kualitas teknis dari model deteksi anomali yang dikembangkan memberikan hasil yang menjanjikan, memberikan bukti empiris yang kuat untuk pergeseran paradigma pengawasan. Kemampuan diskriminasi model disimpulkan melalui nilai ROC-AUC yang mendukung efektivitas deteksi anomali, melampaui ambang batas 0,99. Interpretasi dari skor ini adalah bahwa model memiliki reliabilitas yang sangat tinggi dalam memprioritaskan entitas yang dicurigai melakukan penghindaran pajak, dengan tingkat kesalahan klasifikasi yang minimal. Hal ini secara langsung mendukung prinsip DDDM, di mana keputusan alokasi sumber daya pemeriksaan tidak lagi didasarkan pada intuisi subjektif, melainkan pada skor risiko yang terukur, objektif, dan konsisten.

Keandalan struktural kluster anomali divalidasi oleh *Silhouette Score* 0,94. Skor yang sangat tinggi ini menegaskan bahwa pola anomali yang diidentifikasi oleh model *Autoencoder*, yang menggunakan rasio ETR, CETR, dan DER, secara signifikan berbeda dan kohesif dari pola transaksi normal. Temuan ini secara teoretis mengukuhkan validitas penerapan pendekatan *unsupervised learning* dalam mengidentifikasi pola penyimpangan yang bersifat *hidden* dan dinamis, sebagaimana ditekankan dalam Teori Deteksi Anomali.

Pemilihan model *Autoencoder* dalam penelitian ini didasarkan pada pertimbangan teknis terkait kompleksitas dan dimensi data keuangan. Sebagai perbandingan, terdapat beberapa metode pendeteksi anomali berbasis *unsupervised learning* lain yang lazim digunakan, yaitu *Isolation Forest*, *Local Outlier Factor* (LOF), *One-Class Support Vector Machine* (OC-SVM), dan DBSCAN. Algoritma *Isolation Forest* merupakan metode berbasis pohon keputusan yang efisien. Akan tetapi, metode ini memiliki keterbatasan dalam menangkap interaksi non-linear yang sangat kompleks antar variabel keuangan secara bersamaan. Sementara itu, *Local Outlier Factor* (LOF) yang beroperasi berbasis kepadatan data (*density-based*) cenderung memakan waktu komputasi

yang tinggi pada *dataset* berskala besar, serta kinerjanya sangat sensitif terhadap penentuan parameter *ketetapan* awal. Metode *One-Class SVM* (OC-SVM) pada dasarnya tangguh dalam memisahkan batas data pada ruang berdimensi tinggi. Kendala utama algoritma ini terletak pada aspek *skalabilitas* ketika dihadapkan pada volume data yang masif, karena kompleksitas komputasinya akan meningkat secara kuadratik dan proses penyesuaian *hyperparameter* menjadi sangat sulit tanpa adanya data berlabel. Di sisi lain, algoritma *clustering* tradisional seperti DBSCAN sering kali mengalami kesulitan dalam menangani *dataset* pelaporan keuangan yang memiliki tingkat kepadatan sangat bervariasi dan rentan terhadap kutukan dimensi (*curse of dimensionality*). Berbeda dengan metode alternatif tersebut, arsitektur *Autoencoder* yang berakar pada jaringan saraf tiruan (*deep learning*) memiliki keunggulan komparatif dalam mempelajari representasi fitur secara mendalam. Model ini mampu mereduksi dimensi data ke dalam ruang laten dan merekonstruksinya kembali, sehingga anomali dapat dideteksi secara lebih adaptif dan presisi melalui tingkat kesalahan rekonstruksi (*reconstruction error*). Kapabilitas ini menjadikan *Autoencoder* jauh lebih terukur (*scalable*) dan tangguh dalam memetakan pola penghindaran pajak terselubung dibandingkan algoritma pembandingan lainnya.

Capaian ini menempatkan *TaxSense* pada posisi yang lebih unggul dibandingkan penelitian yang menggunakan pendekatan terawasi (*supervised learning*). Penelitian terdahulu oleh Rosid (2023) dengan ANN (akurasi 92%) dan Beiggi & Ayneband (2024) dengan SVM (akurasi 83,79%) masih bergantung pada data berlabel (*ground truth*), yang dalam konteks penghindaran pajak cenderung sulit diakses, bias, atau ketinggalan zaman. Keunggulan model *unsupervised* terletak pada kemampuan adaptasinya yang inheren terhadap skema penghindaran pajak baru yang belum terlabeli. Selain itu, nilai *Precision Score* 0,88 yang tinggi sangat krusial dalam konteks praktis karena memastikan minimalnya *false positive*, sehingga meminimalkan intervensi audit yang tidak perlu pada wajib pajak yang patuh dan mengalokasikan sumber daya pemeriksaan yang terbatas secara efisien.

Kontribusi Praktis Berdasarkan Validasi Kualitatif (UAT dan ISO 9241-11)

Kehadiran prototipe *TaxSense* terbukti sangat relevan dalam mengatasi kendala operasional di lapangan, sebagaimana ditegaskan oleh hasil validasi kualitatif melalui wawancara mendalam bersama para praktisi perpajakan dan akademisi. Dalam diskusi tersebut, para ahli secara konsisten menyoroti bahwa hambatan pengawasan utama saat ini adalah besarnya volume data yang tidak mampu lagi ditangani secara manual, yang situasinya semakin diperburuk oleh ekstremnya rasio beban kerja para petugas pajak.

Berdasarkan evaluasi *User Acceptance Test* (UAT) yang merujuk pada standar ISO 9241-11, sistem ini terbukti memberikan kontribusi praktis yang nyata melalui tiga dimensi utama. Pertama, dari segi efektivitas, model AI yang dikembangkan berhasil menerjemahkan data keuangan yang rumit menjadi indikator risiko yang jelas dan akurat. Pencapaian ini merupakan bentuk pemenuhan langsung terhadap tujuan utama penelitian, sekaligus membuktikan keberhasilan penerapan Teori Deteksi Anomali.

Kedua, dari sudut pandang efisiensi, *TaxSense* memberikan dampak yang paling signifikan. Proses analisis yang sebelumnya memakan banyak waktu kini dapat diselesaikan secara instan atau *real-time*. Peningkatan kecepatan ini secara fundamental menjawab persoalan ketimpangan beban kerja, mengingat saat ini satu orang petugas sering kali harus mengawasi hingga 10.000 wajib pajak. Selain itu, efisiensi ini juga menguatkan temuan Shakil & Tasnia (2022) yang menyatakan bahwa pemanfaatan AI mampu mendongkrak kinerja administrasi perpajakan.

Ketiga, terkait kepuasan pengguna, desain antarmuka *TaxSense* yang intuitif serta kemampuannya menyajikan skor risiko yang terukur berhasil menciptakan tingkat kepuasan yang tinggi. Peningkatan kualitas interaksi (*usability*) ini sangat krusial guna mendorong tingkat adopsi sistem oleh otoritas pajak, sebuah kondisi yang selaras dengan prinsip-prinsip dalam *User Acceptance Test Theory*.

Sebagai pelengkap, ketangguhan sistem ini juga dibuktikan melalui serangkaian pengujian kualitas non-fungsional yang mencakup *reliability*, *stress*, dan *volume testing*. Hasilnya mengindikasikan bahwa sistem beroperasi dengan sangat stabil, mampu mempertahankan waktu respons di angka 150 ms, serta menunjukkan kinerja yang baik dengan skor 97 pada volume testing.

Walaupun pengujian *stress testing* menunjukkan adanya sedikit degradasi performa (tingkat *error* 18,20%) ketika sistem dihantam beban ekstrem, sistem tetap lolos uji karena mampu terus beroperasi tanpa mengalami kegagalan total. Temuan ini menjadi bukti kuat bahwa *TaxSense* memiliki ketahanan yang memadai untuk segera diimplementasikan di lingkungan produksi nyata.

Kontribusi Ilmiah dan Pengembangan Ilmu Pengetahuan

Secara ilmiah, penelitian ini memberikan kontribusi yang substansial pada diskursus Sistem Informasi Akuntansi (SIA). Penelitian ini tidak hanya memvalidasi sebuah algoritma di atas kertas, tetapi juga menyajikan implementasi model *unsupervised deep clustering* yang terintegrasi secara fungsional ke dalam aplikasi web yang tervalidasi oleh pengguna akhir.

Kontribusi utamanya adalah menyediakan landasan empiris bagi adopsi teknologi mutakhir pada sektor keuangan dan administrasi perpajakan di Indonesia. Model *TaxSense* secara eksplisit menjawab celah penelitian yang diidentifikasi oleh Adamov (2019) dan Rahayu (2021) terkait kebutuhan implementasi *machine learning* dan penyelesaian kendala infrastruktur data di administrasi perpajakan. Dengan mewujudkan sistem yang responsif, terukur, dan mampu mengidentifikasi anomali, penelitian ini membuka peluang riset lanjutan yang menjanjikan, terutama dalam hal pemanfaatan data Big Data yang lebih beragam (misalnya data non-keuangan atau data media sosial) untuk memperkuat basis DDDM dalam pengawasan fiskal. Pada akhirnya, inovasi ini berpotensi besar untuk mengurangi kerugian negara akibat praktik penghindaran pajak yang kompleks, serta mendorong terwujudnya *data-driven organization* di lingkungan Direktorat Jenderal Pajak.

Keterbatasan Penelitian

Meskipun sistem *TaxSense* yang dikembangkan dalam penelitian ini menunjukkan potensi besar untuk mendeteksi anomali perpajakan secara efektif, perlu diakui beberapa keterbatasan yang melingkupinya. Pertama, lingkup variabel keuangan yang dianalisis dalam penelitian ini masih terbatas pada parameter *Effective Tax Rate* (ETR), *Cash Effective Tax Rate* (CETR), *Debt to Equity Ratio* (DER), dan *Profit Before Tax* (PBT). Meskipun variabel tersebut teoretis relevan, penggunaan rasio keuangan yang lebih bervariasi dapat memberikan dimensi analisis yang lebih kaya.

Kedua, meskipun penelitian ini telah menyajikan justifikasi teoretis terkait pemilihan *Autoencoder*, analisis empiris yang membandingkan metrik kinerjanya secara berdampingan (*head-to-head*) dengan algoritma pendeteksi anomali lain seperti *Isolation Forest*, LOF, atau *One-Class SVM* belum dilakukan dan dapat menjadi peluang riset selanjutnya.

Ketiga, dari segi cakupan data, algoritma kecerdasan buatan dalam riset ini dilatih menggunakan *dataset* yang dikhususkan pada entitas korporasi berstatus perusahaan terbuka dengan kriteria inklusi spesifik. Oleh sebab itu, efektivitas sistem dalam mengenali pola anomali pada sektor industri dengan skala menengah atau perusahaan tertutup yang belum melantai di bursa perlu dievaluasi lebih lanjut.

Terakhir, prototipe *TaxSense* saat ini baru melalui tahap uji fungsionalitas dan *User Acceptance Test* di lingkungan eksperimen. Validasi secara empiris dengan mengujicobakan sistem secara langsung pada lingkungan operasional riil di Direktorat Jenderal Pajak, yang terhubung dengan infrastruktur *big data* nasional, merupakan langkah yang krusial namun belum tercakup dalam studi ini.

KESIMPULAN

Penelitian ini telah berhasil memenuhi secara detail ketiga tujuan spesifik yang diusung, melalui pengembangan, implementasi, dan validasi sistem *TaxSense* sebagai solusi untuk deteksi penghindaran pajak. Tujuan utama penelitian ini, yaitu memodelkan dan mengembangkan kecerdasan buatan berbasis *machine learning* untuk mendeteksi kecurangan perpajakan secara otomatis, telah tercapai melalui pembangunan Model *Deep Clustering* berbasis *Autoencoder*. Model ini dilatih secara ekstensif menggunakan 10.052 entri data laporan keuangan historis dari

perusahaan terbuka yang terdaftar di Bursa Efek Indonesia (BEI). Keunggulan fundamental model ini terletak pada penggunaan pendekatan *unsupervised learning*, yang memungkinkan sistem untuk secara adaptif mengidentifikasi pola anomali tersembunyi tanpa bergantung pada data berlabel konvensional yang seringkali bias atau ketinggalan zaman.

Kinerja teknis model *Autoencoder* ini terkonfirmasi menghasilkan performa yang kompetitif, didukung dengan metrik evaluasi yang tinggi. Nilai ROC-AUC yang mencapai 0,99 secara tegas menunjukkan kemampuan diskriminasi yang tinggi dalam membedakan secara akurat antara data normal dan anomali, memberikan reliabilitas yang baik dalam pemetaan risiko. Lebih lanjut, struktur klaster anomali yang terbentuk divalidasi oleh *Silhouette Score* 0,94, yang menegaskan kohesivitas klaster yang sangat baik, sementara *Precision Score* 0,88 memastikan rendahnya risiko *false positive*. Akurasi tinggi ini sangat krusial untuk menjaga efisiensi audit, sehingga sumber daya pemeriksaan dapat dialokasikan hanya kepada entitas yang benar-benar berisiko tinggi.

Tujuan kedua penelitian, yaitu mengimplementasikan model algoritma tersebut ke dalam sebuah purwarupa web (*website prototype*) guna memfasilitasi otoritas pajak dalam pengawasan *real-time*, juga berhasil diwujudkan. Model AI yang telah divalidasi performanya tersebut berhasil diintegrasikan ke dalam purwarupa web *TaxSense*, sebuah aplikasi fungsional yang dirancang khusus untuk mendukung proses pengawasan perpajakan secara instan.

Integrasi model ini secara signifikan memungkinkan konversi data keuangan perusahaan yang rumit menjadi *output* yang cepat dan terukur, berupa Skor Indikasi Penghindaran Pajak. Mekanisme ini didukung oleh penetapan ambang batas deteksi anomali pada persentil ke-90 dari *reconstruction error*, yang secara otomatis melabeli entitas berisiko tinggi dan menyediakan *ground truth* berbasis data untuk tindakan pengawasan. Implementasi ini pada akhirnya memberikan dukungan penuh terhadap prinsip pengambilan keputusan berbasis data (*Data-Driven Decision Making/DDDM*) dalam lingkungan fiskal.

Tujuan ketiga dan terakhir, yaitu mengevaluasi kualitas sistem secara keseluruhan melalui serangkaian pengujian teknis yang terukur guna memastikan kelayakannya, telah divalidasi secara komprehensif. Secara non-fungsional, sistem *TaxSense* menunjukkan ketahanan (*resilience*) dan stabilitas yang tinggi. Hasil *Reliability Testing* mencatat waktu respons rata-rata 150 ms dengan tingkat kesalahan 0.00%, dan *Volume Testing* (diukur oleh *Google PageSpeed Insights*) menghasilkan skor performa unggul 97, yang menandakan kapabilitas aplikasi dalam menangani volume akses tinggi. Meskipun *Stress Testing* menunjukkan degradasi performa pada beban ekstrem (tingkat error 18,20%), sistem tetap lolos uji dengan beroperasi tanpa kegagalan total, membuktikan ketahanan yang memadai untuk lingkungan produksi.

Selain pengujian teknis, validasi kualitatif melalui *User Acceptance Test* (UAT), yang dilakukan dengan praktisi dan akademisi berdasarkan standar ISO 9241-11, mengonfirmasi penerimaan sistem oleh pengguna akhir. UAT menegaskan bahwa sistem unggul dalam tiga aspek utama: Efektivitas (akurasi tinggi dalam identifikasi anomali), Efisiensi (peningkatan kecepatan analisis dari manual menjadi *real-time*), dan Kepuasan Pengguna (mencapai skor positif 8 hingga 10). Secara ringkas, penelitian ini berhasil mengintegrasikan teknologi *machine learning* dan *unsupervised learning* ke dalam sistem informasi akuntansi (*Accounting Information System/AIS*) pengawasan perpajakan. *TaxSense* terbukti menjadi solusi yang responsif, objektif, dan konsisten, mampu mengatasi tantangan pengawasan manual yang disebabkan oleh kompleksitas praktik penghindaran pajak di Indonesia dan volume data yang masif.

Keberhasilan implementasi model ini membuka berbagai peluang strategis bagi arah penelitian selanjutnya. Pengembangan sistem di masa depan dapat difokuskan pada integrasi data non-keuangan, seperti sentimen berita perusahaan dan analisis teks dokumen perpajakan, melalui pemanfaatan teknologi *Natural Language Processing* (NLP). Selain itu, diperlukan studi komparatif secara empiris untuk mengevaluasi performa model *Autoencoder* secara berdampingan dengan algoritma deteksi anomali lainnya, yaitu *Isolation Forest*, *Local Outlier Factor* (LOF), dan *One-Class SVM* (OC-SVM). Pada skala operasional yang lebih luas, riset masa depan diharapkan mampu merintis pengembangan sistem pengawasan perpajakan nasional berbasis *big data* yang dapat terintegrasi secara langsung dengan sistem *Compliance Risk Management* (CRM) milik Direktorat Jenderal Pajak. Sinergi teknologi ini diharapkan dapat menciptakan ekosistem pengawasan fiskal yang lebih komprehensif, presisi, dan mutakhir.

REFERENSI

- Adamov, A. Z. (2019). Machine Learning and Advanced Analytics in Tax Fraud Detection. *2019 IEEE 13th International Conference on Application of Information and Communication Technologies (AICT)*, 1–5. <https://doi.org/10.1109/AICT47866.2019.8981758>
- Alstadsaeter, A., Johannesen, N., Herry, S. L. guern, & Zucman, G. (2021). *Tax Evasion and Tax Avoidance **.
- Asmarani, N. G. C. (2020, Februari 3). *Apa Itu CRM?* <https://news.ddtc.co.id/literasi/kamus/18714/apa-itu-crm>
- Battaglini, M., Guiso, L., Lacava, C., Miller, D. L., & Patacchini, E. (2024). Refining public policies with machine learning: The case of tax auditing. *Journal of Econometrics*, 105847. <https://doi.org/10.1016/j.jeconom.2024.105847>
- Beiggi, H. R., & Ayneband, M. (2024). Detecting Tax Evasion of Legal Entities Using Artificial Intelligence. *Journal of Tax Research*, 34(60), 199–213. <https://doi.org/10.61186/taxjournal.34.60.199>
- Bhattacharyya, A., & Imam, T. (2024). Mandated CSR spending and Tax aggressiveness: A machine learning-driven analysis. *Journal of Cleaner Production*, 452, 142140. <https://doi.org/10.1016/j.jclepro.2024.142140>
- Brynjolfsson, E., & McElheran, K. (2016). Data in Action: Data-Driven Decision Making in U.S. Manufacturing. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2722502>
- Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection. *ACM Computing Surveys*, 41(3), 1–58. <https://doi.org/10.1145/1541880.1541882>
- Devi, Y., Saefurrohman, G. U., Rosilawati, W., Utamie, Z. R., & Nurhayati, N. (2022). Analisis Penyebab Penghindaran Pajak (Tax Avoidance) Dalam Laporan Keuangan Pada Perusahaan Yang Terdaftar di BEI Tahun 2016-2019. *Jurnal Akuntansi dan Pajak*, 22(2), 622. <https://doi.org/10.29040/jap.v22i2.3920>
- Fawcett, T., & Provost, F. (2013). *Data Science for Business*.
- Guenther, D. A., Peterson, K., Searcy, J., & Williams, B. M. (2023). How Useful Are Tax Disclosures in Predicting Effective Tax Rates? A Machine Learning Approach. *The Accounting Review*, 98(5), 297–322. <https://doi.org/10.2308/TAR-2021-0398>
- Hall, J. A. (2016). *Accounting Information Systems* (9 ed.). Cengage Learning Asia Pte Limited.
- Hambling, B., & van Goethem, P. (2013). *User Acceptance Testing: A Step-by-step Guide*. BCS. <https://books.google.co.id/books?id=NzDkkgEACAAJ>
- Hodge, V. J., & Austin, J. (2004). A Survey of Outlier Detection Methodologies. *Artificial Intelligence Review*, 22(2), 85–126. <https://doi.org/10.1007/s10462-004-4304-y>
- Iqbalsah, R. (2023). Machine learning: Classifying taxpayer's supervising zone based on the street address using Natural Language Processing algorithm. *Scientax*, 4(2), 233–242. <https://doi.org/10.52869/st.v4i2.486>
- ISO. (2018). *ISO 9241-11:2018(en), Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*. <https://www.iso.org/obp/ui/en/#iso:std:iso:9241:-11:ed-2:v1:en>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Lenz, H. (2020). Aggressive Tax Avoidance by Managers of Multinational Companies as a Violation of Their Moral Duty to Obey the Law: A Kantian Rationale. *Journal of Business Ethics*, 165(4), 681–697. <https://doi.org/10.1007/s10551-018-4087-8>
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill. <https://books.google.co.id/books?id=EoYBngEACAAJ>
- OECD. (2016). *BEPS Project Explanatory Statement*. OECD. <https://doi.org/10.1787/9789264263437-en>
- Peppers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3), 45–77. <https://doi.org/10.2753/MIS0742-1222240302>

- Rahayu, S. K. (2021). Utilization of Artificial Intelligence in Tax Audit in Indonesia. Dalam *MANAGEMENT AND ACCOUNTING REVIEW* (Vol. 20).
- Romney, M. B., & Steinbart, P. J. (2015). *Accounting Information Systems* (13 ed.). Pearson.
- Rosid, A. (2023). Artificial Neural Networks for predicting taxpaying behaviour of Indonesian firms. *Scientax*, 4(2), 174–204. <https://doi.org/10.52869/st.v4i2.526>
- Russell, S. J., Russell, S., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach*. Pearson. <https://books.google.co.id/books?id=koFptAEACAAJ>
- Sampa, A. W. (2023). *Development Of A Prediction Model For Tax Assessments Using Data Mining And Machine Learning Tools*. <https://dspace.unza.zm/handle/123456789/8800>
- Setyaningsih, W., & Dwi Syamsiah, T. (2024). Tata Kelola Perusahaan dan Penghindaran Pajak pada Perusahaan Multinasional dengan Moderasi Rasio Cakupan Pemeriksaan Pajak. *Jurnal Inovasi Pajak Indonesia*, 1(2), 68–83. <https://doi.org/10.69725/jipi.v1i2.111>
- Shakil, M. H., & Tasnia, M. (2022). Artificial Intelligence and Tax Administration in Asia and the Pacific. Dalam *Taxation in the Digital Economy* (hlm. 45–55). Routledge. <https://doi.org/10.4324/9781003196020-4>
- Shmueli, G., Bruce, P. C., Yahav, I., Patel, N. R., & Lichtendahl, K. C. (2017). *Data Mining for Business Analytics: Concepts, Techniques, and Applications in R*. Wiley. <https://books.google.com/gi/books?id=ETwuDwAAQBAJ>
- Swandi, E. D., & Prasetyo, A. (2024). META ANALISIS DETERMINAN PENGHINDARAN PAJAK. *Jurnal Akuntansi*, 13(1), 44–55. <https://doi.org/10.46806/ja.v13i1.1057>
- Tax Justice Network. (2020, November). *The State of Tax Justice 2020: Tax Justice in the time of COVID-19*. <https://taxjustice.net/reports/the-state-of-tax-justice-2020/>
- Wibowo, B. D. S. (2022). XBRL Open Information Model for Risk Based Tax Audit using Machine Learning. *International Journal of Informatics, Information System and Computer Engineering (INJIISCOM)*, 3(1), 21–46. <https://doi.org/10.34010/injiiscom.v3i1.6891>
- Wilkinson, J. W., Michael J. Cerullo, Vasant Raval, & Bernard Wong-On-Wing. (1999). *Accounting Information Systems: Essential Concepts and Applications* (4 ed.). Wiley.
- Yalamati, S. (2023). Identify fraud detection in corporate tax using Artificial Intelligence advancements Solutions Architect. *International Journal of Machine Learning for Sustainable Development*. <https://ijsdcs.com/index.php/IJMLSD/article/view/468/188>
- Zilli, C. A., Bastos, L. C., & Da Silva, L. R. (2024). Machine learning models in mass appraisal for property tax purposes: a systematic mapping study. *Aestim*, 84, 31–52. <https://doi.org/10.36253/aestim-15792>