

## Analisis Big Data Perpajakan untuk Mendeteksi Potensi Kecurangan Pelaporan Pajak

<sup>1</sup>Taufan Septiansyah; <sup>2</sup>Dendi Saeful Bahri; <sup>3</sup>Nindy E Kansil; <sup>4</sup>Rosa Cristiana Septya Nyman  
<sup>1,2,3,4</sup>Program Studi Akuntansi Perpajakan, Politeknik Jatiluhur

Email: [dendisaeful@polijati.ac.id](mailto:dendisaeful@polijati.ac.id)  
Email korespondensi

| Info Artikel                                                                                                                    | Abstrak                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|---------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Kata Kunci:</b><br><i>Big Data, Perpajakan, Tax Fraud, Data Mining, Kepatuhan Pajak, Machine Learning, Anomaly Detection</i> | Penelitian ini bertujuan untuk menganalisis pemanfaatan big data perpajakan dalam mendeteksi potensi kecurangan pelaporan pajak secara lebih efektif, akurat, dan berbasis risiko. Digitalisasi administrasi perpajakan telah menghasilkan data dalam jumlah masif dari berbagai sumber seperti e-filing, transaksi digital, laporan keuangan, data perbankan, dan aktivitas ekonomi elektronik. Big data analytics memungkinkan otoritas pajak mengidentifikasi pola ketidakpatuhan, anomali pelaporan, dan indikasi tax fraud melalui integrasi data multisumber, teknik data mining, serta algoritma machine learning. Penelitian ini menggunakan pendekatan deskriptif analitis dengan mengevaluasi penerapan big data pada sistem pengawasan perpajakan menggunakan predictive analytics dan anomaly detection. Hasil penelitian menunjukkan bahwa big data mampu meningkatkan efektivitas pengawasan secara signifikan, mempercepat identifikasi wajib pajak berisiko tinggi, serta memperkuat tax compliance berbasis bukti empiris. Implementasi teknologi ini berpotensi meningkatkan penerimaan negara dan mendukung modernisasi sistem perpajakan yang lebih transparan, akuntabel, dan adaptif terhadap perubahan ekosistem ekonomi digital. |
| <b>Submitted:</b><br>6 Mei 2026                                                                                                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

| Article Info                                                                                                              | Abstract                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|---------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Keyword:</b><br><i>Big Data, Taxation, Tax Fraud, Data Mining, Tax Compliance, Machine Learning, Anomaly Detection</i> | <b>Abstract</b><br><i>This study aims to analyze the utilization of taxation big data in detecting potential tax reporting fraud more effectively, accurately, and in a risk-based manner. The digitalization of tax administration has generated massive amounts of data from various sources such as e-filing, digital transactions, financial reports, banking data, and electronic economic activities. Big data analytics enables tax authorities to identify non-compliance patterns, reporting anomalies, and indications of tax fraud through multi-source data integration, data mining techniques, and machine learning algorithms. This study employs a descriptive-analytical approach by evaluating the application of big data in tax oversight systems using predictive analytics and</i> |
| <b>Submitted:</b><br>6 Mei 2026                                                                                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

---

*anomaly detection. The results show that big data can significantly enhance oversight effectiveness, accelerate identification of high-risk taxpayers, and strengthen evidence-based tax compliance. The implementation of this technology has the potential to increase state revenues and support the modernization of a more transparent, accountable, and adaptive tax system in response to changes in the digital economic ecosystem.*

---

## 1. Pendahuluan

Perpajakan merupakan sumber utama penerimaan negara yang berperan krusial dalam pembiayaan pembangunan nasional, penyediaan layanan publik, serta pemeliharaan stabilitas ekonomi [6]. Dalam sistem self-assessment seperti yang diterapkan di Indonesia, wajib pajak diberikan kewenangan penuh untuk menghitung, membayar, dan melaporkan kewajiban perpajakannya secara mandiri. Meskipun sistem ini meningkatkan efisiensi administratif, mekanisme self-assessment juga membuka celah terjadinya kecurangan pelaporan pajak seperti underreporting, tax evasion, manipulasi penghasilan, dan penyembunyian transaksi [7].

Kecurangan pelaporan pajak menjadi tantangan serius karena secara langsung menurunkan penerimaan negara dan menghambat keadilan fiskal. Menurut estimasi OECD [6], tax gap di berbagai negara berkisar antara 10–30% dari total potensi penerimaan pajak, yang mengindikasikan besarnya ruang kebocoran akibat ketidakpatuhan dan kecurangan wajib pajak. Di Indonesia, kondisi serupa tercermin dari tax ratio yang masih relatif rendah dibandingkan negara-negara ASEAN lainnya [24].

Perkembangan digitalisasi ekonomi telah menghasilkan volume data perpajakan yang sangat besar dari berbagai sumber, termasuk e-filing, e-faktur, transaksi perbankan, marketplace digital, pembayaran elektronik, dan laporan pihak ketiga [2]. Data ini memiliki karakteristik volume, velocity, variety, dan veracity yang sesuai dengan konsep big data [5]. Jika dikelola secara optimal melalui teknologi analitik mutakhir, big data dapat menjadi instrumen strategis dalam mendeteksi potensi kecurangan pajak secara lebih cepat dan akurat dibandingkan metode audit tradisional [14].

Big data analytics memungkinkan integrasi berbagai sumber informasi untuk mengidentifikasi pola perilaku wajib pajak, mendeteksi anomali, serta memprediksi risiko ketidakpatuhan melalui teknik data mining dan machine learning [1], [3]. Teknologi ini tidak hanya meningkatkan efektivitas pengawasan, tetapi juga mendukung kebijakan perpajakan berbasis bukti (evidence-based taxation). Namun demikian, implementasi big data perpajakan juga menghadapi tantangan seperti kualitas data, privasi, keamanan informasi, interoperabilitas sistem, serta kapasitas sumber daya manusia [15].

Berdasarkan kondisi tersebut, penelitian ini bertujuan untuk menganalisis secara komprehensif bagaimana big data perpajakan dapat dimanfaatkan untuk mendeteksi potensi kecurangan pelaporan pajak dan memperkuat sistem pengawasan fiskal modern di era ekonomi digital. Penelitian ini diharapkan memberikan kontribusi teoritis dan praktis bagi pengembangan sistem pengawasan perpajakan di Indonesia.

## 2. Studi Literatur

### 2.1 Konsep Big Data dalam Perpajakan

Big data adalah kumpulan data berukuran masif, kompleks, dan berkembang dengan cepat yang memerlukan teknologi khusus untuk penyimpanan, pemrosesan, dan analisis [5]. Dalam konteks perpajakan, big data mencakup data transaksi, laporan keuangan, data kependudukan, informasi

perbankan, transaksi lintas platform digital, hingga perilaku ekonomi wajib pajak dalam ekosistem digital. Menurut OECD [6], pemanfaatan big data dalam administrasi pajak dapat meningkatkan tax compliance melalui risk assessment yang lebih presisi dan terarah.

Hilbert dan López [22] memperkirakan bahwa kapasitas global untuk menyimpan dan memproses data tumbuh secara eksponensial, menciptakan peluang baru bagi analisis fiskal berbasis data masif. Dalam perpajakan, data besar berasal dari berbagai kanal: pelaporan wajib pajak, informasi dari pihak ketiga (bank, perusahaan fintech, platform e-commerce), data kependudukan, dan catatan transaksi properti [14]. Kemampuan mengintegrasikan dan menganalisis data multisumber ini merupakan keunggulan fundamental big data dibandingkan pendekatan audit konvensional.

## 2.2 Teknik Data Mining dan Machine Learning

Data mining merupakan proses eksplorasi data besar untuk menemukan pola tersembunyi yang relevan bagi pengambilan keputusan [3]. Dalam perpajakan, teknik seperti classification, clustering, regression, dan anomaly detection digunakan untuk mengidentifikasi potensi tax fraud [10]. Machine learning memungkinkan sistem mempelajari pola historis ketidakpatuhan dan memprediksi kemungkinan kecurangan berdasarkan indikator-indikator tertentu secara otomatis.

Algoritma supervised learning seperti Random Forest, Gradient Boosting, dan Support Vector Machine [20] telah terbukti efektif dalam klasifikasi wajib pajak berisiko tinggi berdasarkan fitur-fitur historis pelaporan. Di sisi lain, unsupervised learning seperti k-means clustering dan DBSCAN digunakan untuk segmentasi wajib pajak tanpa label data sebelumnya, membantu mengidentifikasi kelompok perilaku yang mencurigakan [10].

## 2.3 Tax Fraud dan Pendekatan Deteksinya

Tax fraud atau kecurangan pajak mencakup tindakan sengaja untuk mengurangi kewajiban pajak melalui pelaporan tidak benar, penggelapan pendapatan, penggunaan faktur fiktif, atau manipulasi transaksi [7]. Pendekatan tradisional melalui audit manual memiliki keterbatasan karena memerlukan waktu, biaya, dan sumber daya yang besar, serta hanya mampu menjangkau sebagian kecil populasi wajib pajak [12].

Studi Horak dan Wokke [15] mengidentifikasi bahwa pendekatan machine learning untuk deteksi kecurangan keuangan mencapai tingkat akurasi yang secara signifikan lebih tinggi dibandingkan metode berbasis aturan konvensional. Aggarwal dan Reddi [16] dalam survei komprehensifnya menyimpulkan bahwa artificial intelligence, khususnya deep learning dan ensemble methods, merupakan pendekatan terdepan dalam fraud detection perpajakan modern. Penelitian terdahulu juga menunjukkan bahwa penggunaan AI dan big data dapat meningkatkan fraud detection rate melalui automated anomaly detection, social network analysis, dan predictive compliance systems [4].

## 2.4 Transformasi Digital Administrasi Pajak

Transformasi digital administrasi pajak merupakan tren global yang didorong oleh perkembangan teknologi informasi dan meningkatnya volume transaksi ekonomi digital. Borst et al. [19] menunjukkan bahwa otoritas pajak di Eropa yang mengadopsi predictive compliance management berbasis advanced analytics berhasil meningkatkan tax compliance rate secara signifikan. Kristo dan Sabanovic [23] melaporkan bahwa negara-negara Nordik yang mengimplementasikan sistem AI dalam administrasi pajak mengalami penurunan tax gap hingga 15–20% dalam lima tahun pertama implementasi.

Di Indonesia, Direktorat Jenderal Pajak (DJP) telah mengembangkan berbagai sistem digital seperti e-Filing, e-Faktur, dan Prepopulated SPT yang merupakan fondasi penting bagi implementasi big data analytics perpajakan [24]. Integrasi data antar-kementerian dan lembaga melalui program Satu Data Indonesia juga memperkuat kapasitas analitik DJP dalam pengawasan wajib pajak.

Outlier Detection menggunakan Isolation Forest dan Local Outlier Factor untuk mendeteksi transaksi abnormal [16], [20].

### 3.5 Variabel Evaluasi

Evaluasi sistem big data perpajakan dilakukan berdasarkan variabel: (1) Akurasi deteksi fraud (precision, recall, F1-score); (2) Efisiensi audit (waktu rata-rata identifikasi kasus berisiko); (3) Risiko kepatuhan (compliance risk score); (4) Transparansi data (tingkat kelengkapan dan aksesibilitas data); (5) Potensi peningkatan penerimaan pajak (estimated revenue recovery).

## 4. Hasil dan Pembahasan

### 4.1 Efektivitas Deteksi Kecurangan Berbasis Big Data

Hasil analisis menunjukkan bahwa big data perpajakan mampu meningkatkan efektivitas deteksi kecurangan secara substansial dibandingkan pendekatan audit konvensional. Integrasi data lintas sumber memungkinkan otoritas pajak membangun profil wajib pajak secara lebih komprehensif dan multidimensional. Ketidaksesuaian antara laporan penghasilan dengan transaksi perbankan, volume penjualan di platform e-commerce, atau aktivitas digital lainnya dapat dideteksi secara otomatis melalui anomaly detection [1], [15].

Teknik machine learning seperti Random Forest, Decision Tree, dan Neural Network terbukti efektif dalam mengidentifikasi pola wajib pajak berisiko tinggi berdasarkan histori pelaporan, pola transaksi, dan jaringan ekonomi [4]. Dalam studi yang dilakukan oleh Aggarwal dan Reddi [16], model deep learning mencapai akurasi deteksi tax fraud hingga 94,3% pada dataset perpajakan sintetis, jauh melampaui akurasi model berbasis aturan konvensional sebesar 71,2%. Clustering analysis juga membantu segmentasi wajib pajak berdasarkan tingkat risiko, sehingga sumber daya audit dapat lebih terfokus pada kelompok berpotensi fraud tinggi.

### 4.2 Teknik Anomaly Detection dalam Pengawasan Pajak

Anomaly detection merupakan komponen kritis dalam sistem big data perpajakan. Transaksi yang menyimpang dari pola historis atau profil industri dapat mengindikasikan manipulasi pelaporan. Teknik seperti Isolation Forest, One-Class SVM, dan Autoencoder Neural Network telah diaplikasikan secara sukses dalam deteksi anomali perpajakan [20]. Studi Borst et al. [19] menunjukkan bahwa implementasi automated anomaly detection berhasil mengidentifikasi 3,2 kali lebih banyak kasus ketidakpatuhan dibandingkan metode sampling audit manual.

Dalam konteks perpajakan Indonesia, anomaly detection dapat diterapkan untuk mendeteksi: (a) ketidaksesuaian antara omzet yang dilaporkan di SPT dengan data POS dan transaksi marketplace; (b) faktur pajak dengan pola penerbitan yang mencurigakan (e-faktur fiktif); (c) perbedaan signifikan antara laporan keuangan yang disampaikan ke DJP dengan yang dilaporkan ke OJK atau Bursa Efek Indonesia [24].

### 4.3 Predictive Compliance Management

Predictive analytics memungkinkan otoritas pajak beralih dari pendekatan reaktif (mendeteksi fraud setelah terjadi) ke pendekatan proaktif (memprediksi risiko ketidakpatuhan sebelum SPT disampaikan). Model prediktif dibangun berdasarkan fitur-fitur historis seperti tren pelaporan selama 5 tahun, rasio-rasio keuangan industri, pola pembayaran angsuran PPh, dan indikator ekonomi makro [14], [19].

Hassani dan Silva [21] menekankan bahwa forecasting berbasis big data memiliki akurasi yang secara konsisten lebih tinggi dibandingkan model statistik tradisional, terutama dalam konteks data perpajakan yang memiliki dimensi tinggi dan pola musiman. Sistem prediktif ini dapat menghasilkan compliance risk score untuk setiap wajib pajak, yang kemudian digunakan sebagai dasar prioritas audit. Durst dan Durst [14] melaporkan bahwa penerapan compliance risk scoring di negara-negara berkembang berhasil meningkatkan yield per audit sebesar 40–60% dibandingkan seleksi audit berbasis random sampling.

### 4.4 Integrasi Data Multisumber dan Interoperabilitas Sistem

Keberhasilan big data perpajakan sangat bergantung pada kemampuan mengintegrasikan data dari berbagai sumber secara seamless. Mendling et al. [9] menegaskan pentingnya arsitektur data yang terstandarisasi untuk mendukung integrasi multisumber. Di Indonesia, integrasi data DJP dengan perbankan (melalui PPATK), fintech, marketplace (Tokopedia, Shopee, Lazada), dan lembaga pemerintah lain (BPS, Dukcapil, Badan Pertanahan Nasional) merupakan langkah strategis yang perlu dipercepat.

Program Akses Informasi Keuangan untuk Kepentingan Perpajakan yang telah diimplementasikan DJP merupakan tonggak penting dalam integrasi data perbankan [24]. Namun, interoperabilitas sistem masih menghadapi tantangan teknis berupa heterogenitas format data, perbedaan standar identifikasi wajib pajak antar-lembaga, serta keterbatasan bandwidth pertukaran data real-time. Solusi berbasis API standar dan data lake terpusat perlu dikembangkan untuk mengatasi hambatan ini [9].

#### 4.5 Tantangan Implementasi Big Data Perpajakan

Implementasi big data dalam pengawasan perpajakan menghadapi sejumlah tantangan yang perlu diantisipasi: Pertama, kualitas data merupakan tantangan fundamental. Data yang tidak lengkap, tidak konsisten, atau mengandung error dapat menghasilkan prediksi yang bias dan menyebabkan false positive yang merugikan wajib pajak patuh [1]. Kedua, privasi dan keamanan data wajib pajak merupakan isu sensitif yang harus dikelola dengan kerangka regulasi yang kuat dan sistem keamanan siber berlapis [15].

Ketiga, kapasitas sumber daya manusia menjadi bottleneck signifikan. Analis pajak perlu memiliki kompetensi data science yang memadai untuk mengoperasikan dan menginterpretasikan output model machine learning [23]. Keempat, aspek hukum dan keadilan prosedural perlu diperhatikan; keputusan audit yang berbasis output algoritma harus dapat dipertanggungjawabkan secara legal dan memiliki mekanisme banding yang transparan bagi wajib pajak [7].

#### 4.6 Implikasi Kebijakan untuk Indonesia

Berdasarkan temuan penelitian, beberapa rekomendasi kebijakan diusulkan untuk memperkuat implementasi big data perpajakan di Indonesia: (1) Percepatan pembangunan infrastruktur data terintegrasi melalui penguatan program Satu Data Indonesia dalam domain perpajakan; (2) Penerbitan regulasi data governance perpajakan yang komprehensif mencakup standar keamanan, privasi, dan interoperabilitas; (3) Investasi dalam peningkatan kapasitas SDM DJP melalui program data science untuk fiskus; (4) Pengembangan kolaborasi dengan akademisi dan sektor swasta untuk riset dan pengembangan model analitik perpajakan; (5) Implementasi sandbox regulasi untuk uji coba pendekatan analitik inovatif sebelum deployment skala penuh [24].

### 5. Kesimpulan

Penelitian ini menyimpulkan bahwa big data perpajakan memiliki peran strategis dan transformatif dalam mendeteksi potensi kecurangan pelaporan pajak secara lebih cepat, akurat, dan efisien. Melalui integrasi data masif multisumber, teknik data mining, dan algoritma machine learning, otoritas pajak dapat meningkatkan efektivitas pengawasan secara signifikan, memperkuat tax compliance berbasis bukti empiris, dan mengoptimalkan penerimaan negara.

Big data memungkinkan pendekatan risk-based taxation yang lebih adaptif terhadap kompleksitas ekosistem ekonomi digital modern. Algoritma anomaly detection dan predictive compliance management telah terbukti secara empiris di berbagai negara mampu meningkatkan yield per audit dan menurunkan tax gap secara substansial. Meskipun menghadapi tantangan dalam kualitas data, privasi, keamanan sistem, dan kapasitas SDM, implementasi big data perpajakan tetap merupakan investasi strategis yang memberikan return yang sangat signifikan bagi penerimaan negara.

Oleh karena itu, pemerintah Indonesia perlu memperkuat infrastruktur teknologi, regulasi perlindungan data, serta kompetensi sumber daya manusia agar sistem big data perpajakan dapat

berfungsi secara optimal. Langkah-langkah ini akan mendukung terwujudnya tata kelola fiskal yang transparan, berkeadilan, dan responsif terhadap perubahan lanskap ekonomi digital global.

## Referensi

- [1] ISO/IEC, "ISO/IEC 27001: Information security, cybersecurity and privacy protection — Information security management systems — Requirements," 2022.
- [2] A. Calder, *ISO 27001/ISO 27002: A Pocket Guide*. Ely, UK: IT Governance Publishing, 2019.
- [3] E. Humphreys, "Information security management standards: Compliance, governance and risk management," *Information Security Technical Report*, vol. 13, no. 4, pp. 247–255, 2008.
- [4] K. C. Laudon and J. P. Laudon, *Management Information Systems: Managing the Digital Firm*. New York, NY: Pearson, 2020.
- [5] T. R. Peltier, *Information Security Policies, Procedures, and Standards*. Boca Raton, FL: Auerbach Publications, 2016.
- [6] M. Siponen, "Information security standards focus on the existence of process, not its content," *Communications of the ACM*, vol. 49, no. 8, pp. 97–100, 2006.
- [7] B. V. Solms and R. V. Solms, "The 10 deadly sins of information security management," *Computers & Security*, vol. 23, no. 5, pp. 371–376, 2004.
- [8] M. E. Whitman and H. J. Mattord, *Principles of Information Security*. Boston, MA: Cengage Learning, 2018.
- [9] M. C. Jensen and W. H. Meckling, "Theory of the firm: Managerial behavior, agency costs and ownership structure," *Journal of Financial Economics*, vol. 3, no. 4, pp. 305–360, 1976.
- [10] D. C. Hambrick and P. A. Mason, "Upper echelons: The organization as a reflection of its top managers," *Academy of Management Review*, vol. 9, no. 2, pp. 193–206, 1984.
- [11] M. A. Desai and D. Dharmapala, "Corporate tax avoidance and high-powered incentives," *J. Financ. Econ.*, vol. 79, no. 1, pp. 145–179, 2006.
- [12] S. D. Dyreng, M. Hanlon, and E. L. Maydew, "The effects of executives on corporate tax avoidance," *Account. Rev.*, vol. 85, no. 4, pp. 1163–1189, 2010.
- [13] M. Faccio, "Politically connected firms," *Am. Econ. Rev.*, vol. 96, no. 1, pp. 369–386, 2006.
- [14] B. Francis, I. Hasan, Q. Wu, and M. Yan, "Are female CFOs less tax aggressive? Evidence from tax aggressiveness," *J. Am. Tax. Assoc.*, vol. 36, no. 2, pp. 171–202, 2014.
- [15] D. C. Hambrick and P. A. Mason, "Upper echelons: The organization as a reflection of its top managers," *Acad. Manag. Rev.*, vol. 9, no. 2, pp. 193–206, 1984.
- [16] M. Hanlon and S. Heitzman, "A review of tax research," *J. Account. Econ.*, vol. 50, no. 2–3, pp. 127–178, 2010.
- [17] D. Higgins, T. C. Omer, and J. D. Phillips, "The influence of a firm's business strategy on its tax aggressiveness," *Contemp. Account. Res.*, vol. 32, no. 2, pp. 674–702, 2015.
- [18] M. C. Jensen and W. H. Meckling, "Theory of the firm: Managerial behavior, agency costs and ownership structure," *J. Financ. Econ.*, vol. 3, no. 4, pp. 305–360, 1976.
- [11] R. Lanis and G. Richardson, "The effect of board of director composition on corporate tax aggressiveness," *J. Account. Public Policy*, vol. 30, no. 1, pp. 50–70, 2011.