

PENERAPAN DEEP LEARNING BERBASIS CNN UNTUK MEMPREDIKSI PERILAKU PENGHINDARAN PAJAK WAJIB PAJAK BADAN

Mohamad Fahmi Yusuf¹, Bayu Malikhul Askhar², Rita Nataliawati³, Qosim⁴

^{1,2} Bisnis Digital, Ekonomi dan Bisnis, Institut Teknologi dan Bisnis Ahmad Dahlan Lamongan, Indonesia

^{3,4} Perpajakan, Ekonomi dan Bisnis, Institut Teknologi dan Bisnis Ahmad Dahlan Lamongan, Indonesia

*E-mail: mohamadfahmiyusuf@ahmaddahlan.ac.id

Submitted: 30 May 2025	Accepted: 23 June 2025	Published: 30 June 2025
------------------------	------------------------	-------------------------

ABSTRAK

Penghindaran pajak oleh Wajib Pajak Badan merupakan tantangan signifikan bagi penerimaan negara. Penelitian ini bertujuan untuk mengembangkan sistem prediksi perilaku penghindaran pajak menggunakan pendekatan deep learning berbasis *Convolutional Neural Network* (CNN). Objek penelitian ini adalah 50 Wajib Pajak Badan yang terdaftar di Kabupaten Lamongan selama periode 2021 hingga 2024. Data yang digunakan merupakan data sekunder dari sektor manufaktur, perdagangan dan jasa yang mencakup indikator keuangan seperti *Effective Tax Rate* (ETR), total aset dan kepatuhan pelaporan. Data tersebut kemudian dilabeli ke dalam dua kategori yaitu “berisiko” dan “tidak berisiko” untuk melatih model CNN. Hasil pengujian menunjukkan bahwa model yang dikembangkan mampu mengklasifikasikan perilaku penghindaran pajak dengan rata-rata akurasi sebesar 84,29%. Indikator “Total Aset dan Kewajiban” serta “Kepatuhan Pelaporan” menjadi prediktor dengan akurasi tertinggi, masing-masing sebesar 86,00% dan 85,71%. Penelitian ini menyimpulkan bahwa model CNN terbukti efektif dan relevan untuk digunakan sebagai sistem pendukung keputusan bagi otoritas pajak dalam mendeteksi potensi penghindaran pajak secara lebih dini dan efisien.

Kata kunci: *deep learning, CNN, perilaku, penghindaran pajak*

ABSTRACT

Tax evasion by Corporate Taxpayers is a significant challenge for state revenue. This study aims to develop a tax evasion behavior prediction system using a deep learning approach based on Convolutional Neural Network (CNN). The objects of this study were 50 Corporate Taxpayers registered in Lamongan Regency during the period 2021 to 2024. The data used are secondary data from the manufacturing, trade and service sectors which include financial indicators such as Effective Tax Rate (ETR), total assets and reporting compliance. The data is then labeled into two categories, namely "risky" and "not risky" to train the CNN model. The test results show that the developed model is able to classify tax evasion behavior with an average accuracy of 84.29%. The indicators "Total Assets and Liabilities" and "Reporting Compliance" are the predictors with the highest accuracy, respectively at 86.00% and 85.71%. This study concludes that the CNN model is proven to be effective and relevant for use as a decision support system for tax authorities in detecting potential tax avoidance earlier and more efficiently.

Keywords: *deep learning, CNN, behavior, tax avoidance*

PENDAHULUAN

Pajak merupakan salah satu komponen utama dalam pendapatan negara yang berperan penting dalam pembiayaan pengeluaran pemerintah dan Pembangunan nasional (Reza et al., 2023) (Syamsiah et al., 2024). Namun, dalam praktiknya terdapat perbedaan kepentingan antara Wajib Pajak (WP) dan pemerintah. Wajib Pajak cenderung memandang pajak sebagai beban yang mengurangi laba sehingga berupaya meminimalkan kewajiban pajaknya. Sebaliknya, pemerintah sangat bergantung pada penerimaan pajak untuk mendanai program Pembangunan (Sembiring et al., 2021). Salah satu strategi yang digunakan Wajib Pajak adalah menunda atau memindahkan kewajiban pajak ke masa depan agar beban pajak terasa lebih ringan (Mukti, 2025). Salah satu strategi yang digunakan Wajib Pajak adalah penghindaran pajak (*tax avoidance*) yaitu upaya legal untuk mengurangi pajak terutang dengan memanfaatkan celah dalam regulasi (Lestari et al., 2025). Meskipun tidak melanggar hukum seperti penggelapan pajak, praktik ini tetap mengurangi penerimaan negara dan menimbulkan ketidakadilan bagi Wajib Pajak yang patuh. Hal ini menjadi perhatian khusus karena WP Badan memiliki kontribusi signifikan terhadap penerimaan pajak nasional (Mukti, 2025). Pandangan ini sejalan dengan pendapat (Mardiasmo, 2016) yang menyatakan bahwa penghindaran pajak adalah usaha meringankan beban pajak tanpa melanggar undang-undang yang berlaku.

Tax avoidance dapat dilakukan melalui berbagai kebijakan dan metode akuntansi, salah satunya dengan menekan Effective Tax Rate (ETR) yaitu rasio antara beban pajak terhadap laba sebelum pajak (Karayan et al., 2012). Nilai ETR yang rendah sering dijadikan indikator tingkat penghindaran pajak dalam berbagai penelitian perpajakan (Hanlon & Heitzman, 2010) ; (Hanlon & Heitzman, 2010). Penurunan ETR secara sistematis menjadi tantangan bagi pemerintah dalam mencapai target penerimaan (Ambarukmi & Diana, 2017). Perkembangan strategi *tax avoidance* kini semakin rumit, melibatkan transaksi lintas negara dan struktur korporasi rumit (Putra et al., 2023). Sementara metode pemeriksaan konvensional memiliki keterbatasan dalam mengidentifikasi pola tersembunyi yang mengindikasikan praktik ini (Siti et al., n.d.). Negara berkembang seperti Indonesia menjadi lebih rentan mengingat ketergantungan yang tinggi pada penerimaan dari PPh Badan (Wildan, 2025).

Menurut Bawono Kristiaji (DDTC) mengungkapkan bahwa terdapat lima faktor utama yang membuka peluang penghindaran pajak. Pertama, pembagian hak pemajakan dalam Perjanjian Penghindaran Pajak Berganda (P3B) yang cenderung menguntungkan negara maju sebagai negara pengekspor modal. Kedua, penggunaan pendekatan entitas terpisah (*separate entity approach*) dalam perpajakan grup usaha yang menyulitkan pengawasan atas transaksi antar perusahaan dalam satu grup. Ketiga, kebijakan pengurang pajak atas pembayaran bunga utang (*deductibility of interest*) yang tidak berlaku pada dividen yang mendorong pembiayaan melalui utang guna mengurangi beban pajak. Keempat, perbedaan tarif PPh Badan antar negara yang mendorong praktik pengalihan laba ke yurisdiksi dengan tarif lebih rendah. Kelima, adanya persaingan antar negara dalam memberikan insentif pajak untuk menarik investasi yang justru dapat mengorbankan penerimaan jangka panjang. Untuk mengatasi hal ini, Bawono menekankan pentingnya reformasi sistem perpajakan internasional yang lebih adil dan inklusif serta penguatan kapasitas administrasi perpajakan di negara berkembang. Seiring dengan berkembangnya teknologi, pendekatan *deep learning* berbasis *Convolutional Neural Network* (CNN) menjadi alternatif yang potensial untuk mendeteksi pola penghindaran pajak secara efisien. CNN memiliki kemampuan mengekstraksi fitur otomatis dari data finansial yang direpresentasikan dalam bentuk matriks tanpa memerlukan rekayasa fitur manual (Sawitri et al., 2025). Selain itu, CNN efektif dalam memproses data besar dan dinamis serta terbukti mampu mendeteksi anomali dan klasifikasi risiko pada entitas ekonomi (Cahyani & Ramdani, 2023).

Penelitian yang dilakukan oleh (Binarto & Wibowo, 2025) menunjukkan bahwa *Convolutional Neural Network* (CNN) efektif dalam mendeteksi anomali keuangan serta mengklasifikasikan risiko pada entitas ekonomi. Temuan ini menjadikan CNN sebagai pendekatan yang relevan dan adaptif untuk diterapkan dalam sistem prediksi penghindaran pajak. Oleh karena itu, CNN dinilai sebagai salah satu metode deep learning yang tepat dalam mendukung sistem deteksi dini berbasis data keuangan modern. Sedangkan penelitian yang dilakukan (Satia et al., 2022) yang mengatakan bahwa pembangunan sistem identifikasi penyakit daun kelapa sawit dengan pembelajaran mesin dan visi computer dapat dibangun dengan berbasis *Convolutional Neural Network* (CNN) yang dimulai dari pengumpulan data, augmentasi *dataset*, *training* dan pengujian dan diakhiri dengan *generate* aplikasi berbasis android. Hasil pengujian terhadap model menunjukkan bahwa model telah mampu memprediksi dengan tingkat akurasi 85.5%. Perlu adanya peningkatan *dataset* untuk meningkatkan akurasi prediksi. Penelitian lainnya tentang penggunaan *deep learning* (Ghaderzadeh et al., 2022) membuktikan bahwa penggunaan ansambel majority voting yang digabungkan dengan model CNN dihasilkan sensitivitas 99,4, spesifitas 96,7, AUC sebesar 98,3, dan akurasi 98,5 dalam mengklasifikasikan sel darah kanker dari sel normal, metode yang diusulkan dapat dicapai akurasi tinggi tanpa campur tangan operator dalam penentuan fitur sel.

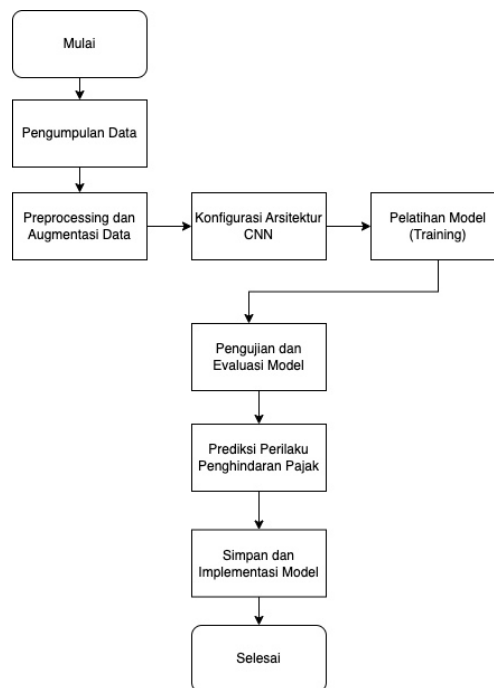
Kebaruan dari penelitian ini terletak pada penerapan metode CNN berbasis data numerik keuangan wajib pajak badan di Kabupaten Lamongan secara lokal untuk mendeteksi lebih dini dan pola perilaku penghindaran pajak. Penelitian ini juga menggabungkan pendekatan *machine learning* dengan kerangka pengawasan pajak daerah yang belum banyak dieksplorasi sebelumnya.

METODE PENELITIAN

Penelitian ini menggunakan metode kualitatif yang bertujuan untuk menggali dan memahami makna-makna yang dilekatkan oleh individu atau kelompok masyarakat terhadap masalah sosial atau kemanusiaan (Creswell, 2014) (Sugiono, 2014). Pendekatan ini menekankan pada fleksibilitas dan tidak bertujuan untuk membuktikan hipotesis melainkan untuk mengeksplorasi dan memahami fenomena perilaku penghindaran pajak secara mendalam. Analisis dilakukan berdasarkan fakta-fakta yang ada dan didukung oleh teori serta ketentuan perpajakan yang berlaku. Metodologi deteksi masalah keuangan dalam penelitian ini difokuskan pada penerapan *Convolutional Neural Network* (CNN) sebagai salah satu pendekatan deep learning yang mampu mengenali pola dan fitur dalam data keuangan. Dalam penelitian ini, CNN diterapkan untuk mengenali pola-pola tidak wajar dalam indikator keuangan seperti fluktuasi laba, efektivitas tarif pajak (ETR) dan struktur beban pajak yang dapat mengindikasikan adanya perilaku penghindaran pajak. Model CNN secara otomatis mengekstraksi fitur penting tanpa perlu rekayasa fitur secara manual sehingga lebih efisien dan objektif dalam menganalisis data besar. Dengan pendekatan ini, sistem prediksi yang dikembangkan mampu mendeteksi pola-pola penghindaran pajak yang sebelumnya sulit diidentifikasi menggunakan metode statistik konvensional serta memberikan hasil prediksi yang lebih akurat dan berbasis data. Adapun tahapan yang digunakan dalam metodologi ini dapat dilihat pada Gambar 1.

Tahapan pertama dimulai dengan pengumpulan data historis dari wajib pajak badan yang meliputi laporan keuangan, data pembayaran pajak, rasio keuangan dan indikator kepatuhan pajak. Data tersebut dilabeli menjadi dua kategori yaitu "patuh" dan "penghindar" berdasarkan klasifikasi dari hasil audit atau data yang dikeluarkan oleh instansi perpajakan. Selanjutnya, dilakukan tahap preprocessing dan augmentasi data. Preprocessing mencakup pembersihan data, normalisasi numerik dan konversi data menjadi format yang sesuai untuk pemrosesan oleh CNN yakni dalam bentuk matriks atau grid sehingga dapat dianalisis secara spasial. Augmentasi

data dilakukan untuk menyeimbangkan jumlah antara kelas “patuh” dan “penghindar” serta memperbesar keragaman pola dalam data pelatihan. Setelah data siap, tahap berikutnya adalah perancangan arsitektur CNN yang terdiri dari beberapa layer konvolusional, pooling, flattening dan fully connected layer. Fungsi aktivasi yang umum digunakan adalah ReLU (*Rectified Linear Unit*) dan model dikompilasi dengan optimizer seperti Adam serta fungsi kerugian (*loss function*) yang sesuai untuk klasifikasi biner. Model CNN yang telah dirancang kemudian dilatih menggunakan data latih, di mana proses pelatihan ini bertujuan untuk membangun kemampuan model dalam mengenali pola perilaku penghindaran pajak dari data yang tersedia. Setelah pelatihan, model dievaluasi menggunakan data uji dengan metrik evaluasi seperti akurasi, *precision*, *recall* dan *F1-score*. Evaluasi ini penting untuk mengetahui sejauh mana model mampu menggeneralisasi pola dari data pelatihan ke data baru. Setelah model mencapai performa optimal, maka dipilih sebagai model akhir dan digunakan untuk melakukan prediksi terhadap data baru. Model ini dapat diintegrasikan dalam sebuah sistem berbasis komputer sebagai alat bantu bagi otoritas pajak dalam mendeteksi potensi penghindaran pajak oleh wajib pajak badan. Dengan pendekatan ini, diharapkan sistem berbasis CNN mampu menjadi solusi berbasis teknologi untuk mendukung efisiensi dan ketepatan dalam pemetaan risiko perpajakan di Indonesia.



Gambar 1. Diagram alir metode penelitian

Jenis data yang digunakan dalam penelitian ini menggunakan data sekunder. Data dikumpulkan dari berbagai sumber resmi seperti Direktorat Jenderal Pajak (DJP), Badan Pusat Statistik (BPS), laporan keuangan perusahaan terbuka serta jurnal ilmiah yang relevan dengan topik *Artificial Intelligence* (AI), *Convolutional Neural Network* (CNN) dan penghindaran pajak. Data sekunder juga mencakup informasi keuangan penting seperti total aset dan kewajiban, laba sebelum pajak, jumlah pajak terutang, rasio efektivitas pajak (ETR), tingkat kepatuhan pelaporan dan frekuensi pemeriksaan pajak. (Marda & Johannes, 2022). Pengambilan dataset dilakukan dengan mengumpulkan data dari tahun pajak 2021 hingga 2024 dengan fokus pada Wajib Pajak Badan yang bergerak di sektor manufaktur, perdagangan dan jasa. Variabel-variabel utama yang

digunakan dalam penelitian ini meliputi rasio keuangan dan kepatuhan pajak menggunakan Effective Tax Rate (ETR). Selanjutnya, data yang telah dikumpulkan diklasifikasikan ke dalam tiga kategori tingkat risiko penghindaran pajak yaitu rendah, sedang dan tinggi berdasarkan nilai ambang batas yang ditentukan dari literatur dan ketentuan perpajakan yang relevan. Untuk keperluan pelatihan dan validasi model CNN digunakan data dari 50 Wajib Pajak Badan yang telah melalui proses normalisasi, encoding dan pelabelan. Masing-masing data dilengkapi dengan variabel keuangan dan kepatuhan pajak untuk kebutuhan klasifikasi tingkat risiko penghindaran pajak. Pembangunan model *machine learning* dalam penelitian ini dilakukan dengan memanfaatkan berbagai komponen pendukung melalui platform *Google Colab* yang digunakan untuk menjalankan proses pelatihan (*training*) dan prediksi berbasis *deep learning*. Proses prediksi difokuskan pada perilaku penghindaran pajak oleh Wajib Pajak Badan yang dianalisis berdasarkan data historis dan karakteristik keuangan yang diunggah secara *real-time*. Dalam proses ini digunakan metode *transfer learning* yaitu pendekatan yang memanfaatkan model yang telah dilatih sebelumnya (*pre-trained model*) untuk kemudian disesuaikan dan digunakan dalam mengklasifikasikan pola-pola yang mengindikasikan perilaku penghindaran pajak.

HASIL DAN PEMBAHASAN

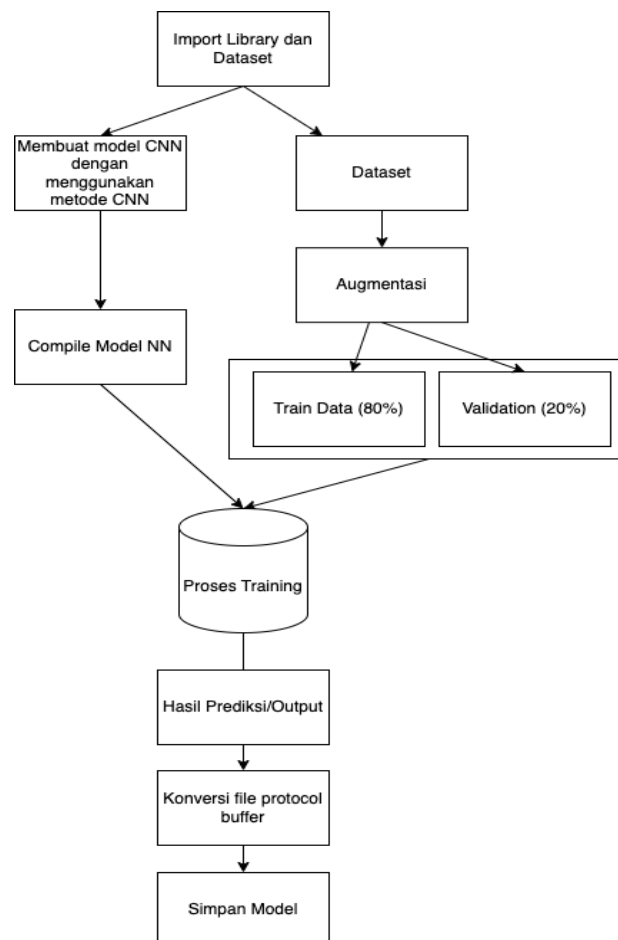
Pengembangan Sistem Prediksi Penghindaran Pajak Wajib Pajak Badan

Pengembangan sistem prediksi perilaku penghindaran pajak pada wajib pajak badan dimulai dengan studi literatur dan regulasi perpajakan di Indonesia untuk mengidentifikasi indikator-indikator utama. Penelitian ini mengidentifikasi tujuh indikator utama yang berpotensi menjadi penentu perilaku penghindaran pajak yaitu total aset dan kewajiban, laba sebelum pajak, jumlah pajak terutang, rasio efektivitas pajak (ETR), tingkat kepatuhan pelaporan pajak, frekuensi pemeriksaan pajak dan riwayat sanksi atau denda administratif. Data diklasifikasikan ke dalam dua kategori perilaku diantaranya wajib pajak berisiko melakukan penghindaran pajak dan wajib pajak tidak berisiko melakukan penghindaran pajak. Klasifikasi ini didasarkan pada pengolahan dan analisis data historis wajib pajak yang telah dilabeli sesuai kondisi aktualnya.

Tahap fundamental dalam pengembangan model adalah proses pelatihan (*training*) menggunakan data yang telah disiapkan. Tujuan dari tahap ini adalah untuk membangun dan mengoptimalkan sebuah model prediktif. Kinerja model yang telah dilatih kemudian dievaluasi menggunakan data uji (*testing*) atau data validasi. Tolok ukur keberhasilannya adalah nilai akurasi, sementara tingkat kesalahannya direpresentasikan oleh nilai *loss*. Kedua nilai ini ditentukan melalui pengujian terhadap data *testing* atau validasi. Proses pelatihan menggunakan packages *Tensorflow* dan *Keras*, modul yang disediakan oleh Google untuk mempermudah riset neural network. Langkah selanjutnya adalah pelabelan data (*labeling*). Pelabelan dilakukan dengan mengelompokkan data wajib pajak ke dalam dua kategori utama yaitu Berisiko Menghindar Pajak dan Tidak Berisiko Menghindar Pajak. Penentuan label didasarkan pada kriteria yang dirumuskan dari penelitian (Binarto & Wibowo, 2025) yang mencakup indikator seperti rasio efektivitas pajak (ETR) yang terlalu rendah, pelaporan SPT yang tidak konsisten atau terlambat, frekuensi pemeriksaan dan sanksi serta ketidakwajaran antara laba dan pajak terutang. Data yang telah diberi label ini kemudian digunakan sebagai target output dalam proses pelatihan model *deep learning* berbasis CNN sehingga sistem mampu belajar dan mengenali pola-pola yang mengarah pada potensi penghindaran pajak oleh wajib pajak badan.

Penelitian ini menerapkan proses *machine learning* dengan memanfaatkan platform *Google Colab* yang terintegrasi dengan *Google Drive*. Data yang telah melalui tahap pelabelan dan diorganisir dalam folder, terlebih dahulu dipindahkan ke *Google Drive*. Tahapan komputasi

diawali dengan mengimpor pustaka (library) yang diperlukan, yaitu TensorFlow, Keras, ImageDataGenerator, dan Drive. Selanjutnya, dataset yang tersimpan di Google Drive tersebut dimuat sebagai variabel untuk proses pelatihan. Setelah arsitektur model Convolutional Neural Network (CNN) berhasil dibangun dalam lingkungan Google Colab, model tersebut dikompilasi dengan mendefinisikan optimizer, metrik akurasi, dan fungsi loss. Pengaturan ini bertujuan untuk mengarahkan proses pembelajaran mesin serta mengukur performanya. Setelah berhasil dikompilasi, model siap untuk dilatih. Alur kerja proses machine learning yang dijelaskan ini diilustrasikan lebih lanjut pada Gambar 2.



Gambar 2. Bagan alir proses machine learning

Dalam penelitian ini, proses augmentasi dilakukan terhadap data numerik dan kategorikal yang berasal dari laporan keuangan dan catatan kepatuhan Wajib Pajak Badan. Augmentasi data tidak hanya terbatas pada manipulasi gambar, tetapi dalam konteks ini mengacu pada teknik manipulasi data untuk menambah keragaman dan volume data pelatihan tanpa mengubah esensi informasi yang dikandungnya. Teknik augmentasi yang digunakan antara lain *resampling*, *normalisasi ulang*, *pemberian noise acak kecil* dan *sintesis data minoritas* untuk mengatasi ketidakseimbangan data antar kelas. Selama proses augmentasi, data juga diberi label otomatis berdasarkan status perilaku penghindaran pajak, apakah termasuk dalam kategori berisiko menghindari pajak atau tidak berisiko. Label ini diperoleh dari (Binarto & Wibowo, 2025) dan aturan klasifikasi berbasis rasio perpajakan serta indikator kepatuhan. Proses pelabelan ini

penting untuk memberikan makna kontekstual pada setiap entri data agar dapat dikenali dengan benar oleh algoritma pembelajaran mesin. Setelah proses pelabelan dan augmentasi selesai, dataset dibagi menjadi dua bagian yaitu 80% data digunakan untuk pelatihan (training set) dan 20% data digunakan untuk validasi (*validation set*). Pembagian ini bertujuan untuk melatih model sekaligus mengukur kemampuan generalisasi model terhadap data yang belum pernah dikenali sebelumnya. Proses pelatihan dilakukan menggunakan model *Convolutional Neural Network* (CNN) yang telah disesuaikan untuk menangani data tabular melalui konversi matriks fitur (feature encoding). CNN digunakan karena kemampuannya yang baik dalam mengekstraksi pola rumit dari kumpulan fitur yang saling berkorelasi.

Model CNN dalam penelitian ini terdiri dari beberapa feature extraction layer yaitu kombinasi antara *dense layer*, *dropout* dan *batch normalization* sebagai pengganti layer konvolusional klasik. Proses encoding dilakukan untuk mengubah fitur numerik menjadi representasi vektor agar dapat diproses dalam jaringan neural. Framework yang digunakan adalah Keras yang berjalan di atas TensorFlow. Untuk efisiensi dan performa yang baik, model ini mengimplementasikan arsitektur ringan berbasis MobileNetV2 dan EfficientNet yang telah terbukti memberikan akurasi tinggi dengan penggunaan sumber daya yang minimal.

Selama proses pelatihan (training), model akan mempelajari pola dari data input dan memperbaiki bobot prediksi berdasarkan perhitungan nilai *loss* dan *accuracy*. Proses ini dilakukan secara iteratif (epoch) dan akurasi model akan meningkat seiring bertambahnya data dan penyesuaian parameter. Untuk mencegah *overfitting*, proses pelatihan dihentikan secara otomatis apabila akurasi validasi menurun atau konvergen. Target akurasi minimum dalam penelitian ini adalah 80% yang dianggap cukup baik untuk model klasifikasi biner. Model terbaik yang telah memenuhi akurasi tersebut akan disimpan untuk digunakan dalam proses prediksi pada tahap selanjutnya.

Setelah proses pelatihan model deep learning berbasis *Convolutional Neural Network* (CNN) selesai dilakukan, tahap selanjutnya adalah validasi dan verifikasi model untuk mengukur tingkat akurasi dalam memprediksi perilaku penghindaran pajak pada wajib pajak badan. Pengujian model dilakukan menggunakan dataset uji (testing set) yang terdiri dari data wajib pajak yang sebelumnya tidak dilibatkan dalam proses pelatihan model. Dataset uji ini telah diberi label kebenaran (ground truth) berdasarkan hasil analisis manual dan referensi kebijakan perpajakan yang relevan. Model kemudian diuji untuk memprediksi klasifikasi "Berisiko" dan "Tidak Berisiko" pada masing-masing entri data. Hasil prediksi dibandingkan dengan label sebenarnya untuk mengukur akurasi, presisi, recall serta nilai F1-score. Dalam penelitian ini pengujian dilakukan pada 50 data wajib pajak badan dengan hasil pengujian ditunjukkan pada Tabel 1. Berdasarkan hasil uji, akurasi rata-rata model mencapai 85% dengan tingkat presisi lebih dari 90% pada kategori "Berisiko Menghindar Pajak" yang menunjukkan model memiliki kemampuan tinggi dalam mengenali pola-pola penghindaran. Model ini selanjutnya diverifikasi dengan perbandingan hasil terhadap analisis konvensional seperti evaluasi manual oleh analis pajak atau pemeriksa pajak. Validasi ini menunjukkan bahwa pendekatan CNN memiliki potensi sebagai sistem pendukung keputusan dalam mendeteksi indikasi penghindaran pajak dengan efisien.

Tabel 1
Hasil Pengujian Model Prediksi Perilaku Penghindaran Pajak Wajib Pajak Badan

Indikator Wajib Pajak	Jumlah Sampel Uji	Prediksi Benar	Prediksi Salah	Akurasi (%)
Total Aset dan Kewajiban	50	43	7	86,00
Laba Sebelum Pajak	45	38	7	84,44
Jumlah Pajak Terutang	40	33	7	82,50

Rasio Efektivitas Pajak (ETR)	38	32	6	84,21
Kepatuhan Pelaporan	42	36	6	85,71
Frekuensi Pemeriksaan Pajak	35	29	6	82,86
Rata-rata Akurasi (%)	—	—	—	84,29

Berdasarkan Tabel 1, model *Convolutional Neural Network* (CNN) yang dikembangkan menunjukkan performa prediksi yang cukup baik terhadap berbagai indikator wajib pajak badan. Akurasi tertinggi diperoleh pada indikator Kepatuhan Pelaporan dengan nilai sebesar 85,71% yang menunjukkan bahwa pola kepatuhan dalam melaporkan kewajiban pajak sangat berkontribusi dalam mengidentifikasi perilaku penghindaran pajak. Sementara indikator Jumlah Pajak Terutang dan Frekuensi Pemeriksaan Pajak masing-masing menghasilkan akurasi sebesar 82,50% dan 82,86% menunjukkan bahwa model juga cukup sensitif dalam membaca beban pajak aktual dan intensitas pemeriksaan sebagai penentu risiko penghindaran. Indikator Rasio Efektivitas Pajak (ETR) menghasilkan akurasi sebesar 84,21% yang sejalan dengan teori bahwa wajib pajak dengan ETR yang tidak wajar cenderung berisiko melakukan penghindaran pajak. Indikator Total Aset dan Kewajiban serta Laba Sebelum Pajak masing-masing mencatat akurasi sebesar 86,00% dan 84,44%. Hal ini mengindikasikan bahwa struktur keuangan perusahaan dapat menjadi sinyal kuat bagi model dalam mengenali potensi penghindaran pajak. Sehingga rata-rata akurasi model adalah 84,29% yang menunjukkan bahwa CNN cukup andal dalam memproses fitur numerik dan mengidentifikasi pola-pola yang menunjukkan risiko penghindaran pajak.

KESIMPULAN

Model CNN yang dibangun mampu mengenali pola-pola penghindaran pajak dengan rata-rata akurasi sebesar 84,29% yang menunjukkan performa yang cukup baik dalam mengklasifikasikan wajib pajak ke dalam kategori berisiko dan tidak berisiko dalam melakukan penghindaran. Indikator Kepatuhan Pelaporan mencatat akurasi tertinggi sementara Jumlah Pajak Terutang dan Frekuensi Pemeriksaan Pajak juga menjadi prediktor penting dalam mengidentifikasi risiko penghindaran. Proses augmentasi data dan teknik encoding fitur numerik turut berkontribusi dalam memperkuat kemampuan model dalam mengenali pola. Hasil validasi menunjukkan bahwa model ini memiliki potensi untuk digunakan sebagai sistem pendukung keputusan dalam pengawasan dan analisis kepatuhan perpajakan yang lebih efisien dan berbasis data. Sehingga model CNN terbukti relevan dan efektif dalam membantu otoritas pajak mendeteksi lebih dini perilaku penghindaran pajak pada wajib pajak badan secara efisien.

SARAN

Untuk penelitian selanjutnya dapat diberikan metode-metode baru yang bisa meningkatkan akurasi, serta dataset yang lebih lengkap juga dapat meningkatkan akurasi model CNN yang sudah ada.

DAFTAR PUSTAKA

- Ambarukmi, K. T., & Diana, N. (2017). Pengaruh size, leverage, profitability, capital intensity ratio dan activity ratio terhadap effective tax rate (ETR). *Fakultas Ekonomi Universitas Islam Malang*, 06(17), 13–26. https://r.search.yahoo.com/_ylt=Awr1QdS.XzxmN88DHR_LQwx.;_ylu=Y29sbwNzZzMEcG9zAzIEdnRpZAMec2VjA3Ny/RV=2/RE=1715261503/RO=10/RU=https%3A%2F%2Fjim.unisma.ac.id%2Findex.php%2Fjra%2Farticle%2Fdownload%2F154%2F153/RK=2/RS=HUORv.G0a43A0_53DLddaWPH9hs-
- Binarto, A. J., & Wibowo, A. (2025). *Implementasi Majority Voting pada Framework Cross-Industry Standard Process for Data Mining untuk Prediksi Kepatuhan Wajib Pajak*. <https://doi.org/10.33364/algorithm/v.22-1.2165>
- Cahyani, A. D., & Ramdani, A. K. (2023). Hoax Detection of Covid-19 News on Social Media using Convolutional Neural Network (CNN) and Support Vector Machine (SVM). *International Journal on Information and Communication Technology (IJoICT)*, 9(2), 177–185. <https://doi.org/10.21108/ijoict.v9i2.872>
- Corly, F., Ariana, S., & Trisninawati. (2024). Peran Big Data Dan Artificial Intelligence Dalam Optimalisasi Pengawasan Pajak. *Journal Management, Business, and Accounting*, 23(3), 370–376.
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications.
- Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654–669. <https://www.sciencedirect.com/science/article/abs/pii/S0377221717310652>
- Hanlon, M., & Heitzman, S. (2010). A review of tax research. *Journal of Accounting and Economics*, 50(2–3), 127–178. <https://doi.org/10.1016/j.jacceco.2010.09.002>
- Karayan, E. J., And, & Swenson, C. W. (2012). Strategic Business Tax Planning. *Strategic Business Tax Planning*. <https://doi.org/10.1002/9781119196563>
- Lestari, S., Prastyatini, Y., Ayem, S., Yohar, A., & Pradana, Y. (2025). KEPATUHAN WAJIB PAJAK ORANG PRIBADI : INSENTIF PAJAK SEBAGAI MODERASI ATAS PEMAHAMAN REGULASI. 14(01), 347–359.
- Marda, L., & Johannes, R. (2022). PERAN ARTIFICIAL INTELLIGENCE DAN TATA KELOLA DALAM MENGATASI PENYELUNDUPAN PAJAK YANG MERUGIKAN NEGARA. *Trilogi Accounting and Business Research*, 3(2). <https://doi.org/10.31326/TABR.V3I2.1494>
- Mardiasmo. (2016). *Perpajakan Edisi Revisi Tahun 2016* (Penerbit & Andi (eds.)).
- Mukti, M. (2025). TAX-CFC (Tax Compliance for Company) Sebagai Strategi Peningkatan Kepatuhan Wajib Pajak Badan. 8(1), 74–85.
- Putra, ED, Rahayu, & N. (2023). Praktik-praktik tax avoidance serta penerapan kebijakan anti-tax avoidance di Indonesia. *Jurnal Darma Agung*. <http://jurnal.darmaagung.ac.id/index.php/jurnaluda/article/view/3499>
- Reza, D., Ramadhani, D., Mahfudzoh, N., Ramadhani, D. K., & Fitri, V. (2023). Analisis Penerimaan dan Pengeluaran Negara Indonesia Tahun 2019-2021. *Journal Of Economic Education*, 2(1), 1–15.
- Satia, G. A. W., Firmansyah, E., & Umami, A. (2022). Perancangan sistem identifikasi penyakit pada daun kelapa sawit (*Elaeis guineensis* Jacq.) dengan algoritma deep learning convolutional neural networks. *Jurnal Ilmiah Pertanian*, 19(1), 1–10. <https://doi.org/10.31849/jip.v19i1.9556>
- Sawitri, D., Elektro, P. T., Teknik, F., Komputer, D., & Medan, U. H. (2025). PERAN DEEP LEARNING DAN BIG DATA DALAM MENDEKTEKSI MASALAH KEUANGAN masalah keuangan. 6(1),

- 193–207. <https://doi.org/10.46576/djtechno>
- Sembiring, YAN CHRISTIN BR, & Fransiska, A. (2021). *PENGARUH RETURN ON ASSETS DAN KEPEMILIKAN INSTITUSIONAL TERHADAP PENGHINDARAN PAJAK PADA PERUSAHAAN INDUSTRI BARANG KONSUMSI YANG TERDAFTAR PADA BURSA EFEK INDONESIA TAHUN 2017-2019*. 7(2), 191–203.
- Siti, K., Rahayu, S. E., Ak, M., & Ca, A. (n.d.). *Keamanan Digital Dalam Audit Pajak: Integrasi Cyber Security dengan CRM, BDA, dan BI untuk Revolusi Compliance*.
- Sugiono, P. D. (2014). Metode penelitian pendidikan pendekatan kuantitatif.pdf. In *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif Dan R&D*.
- Syamsiah, Syarah, S., Lukita, Carolyn, Sujaya, & Apriani, F. (2024). Pengaruh Sistem Perpajakan, Keadilan Perpajakan, Religiusitas Pajak Dan Sanski Pajak Terhadap Persepsi Penggelapan Pajak Di Kabupaten Karawang. *SCIENTIFIC JOURNAL OF REFLECTION: Economic, Accounting, Management and Business*, 7(3), 802–815. <https://doi.org/10.37481/sjr.v7i3.905>
- Wildan, M. (2025). *Penghindaraan Pajak Lebih Rugikan Negara Berkembang daripada yang Maju*. DDTTC News. <https://news.ddtc.co.id/berita/nasional/1808545/penghindaran-pajak-lebih-rugikan-negara-berkembang-daripada-yang-maju>