

BAB I

PENDAHULUAN

1.1 Latar Belakang

Laporan keuangan merupakan keluaran akhir dari proses akuntansi, dan menjadi media utama bagi emiten untuk menyampaikan kondisi keuangannya (Yulianti, 2025). Munculnya isu kecurangan laporan keuangan atau *fraudulent financial statement* menjadi permasalahan yang serius dalam dunia akuntansi dan pasar modal. Kecurangan laporan keuangan atau yang juga bisa disebut dengan *fraudulent financial statmenet* merupakan tindakan yang sengaja dilakukan untuk menyesatkan pihak-pihak yang menggunakan informasi pada laporan tersebut (Manurung & Wirogioto, 2023). Fenomena ini tidak hanya mengurangi kepercayaan investor, tetapi juga mengganggu stabilitas pasar serta mencederai kredibilitas otoritas pengawas.

Data dari *Association of Certified Fraud Examiners* (ACFE) menunjukkan bahwa Indonesia menempati peringkat keempat di Asia Pasifik dalam jumlah kasus kecurangan (*fraud*), mencatat 23 kasus sepanjang tahun 2022. Meskipun proporsi kecurangan laporan keuangan hanya sekitar 9 persen dari seluruh kasus, kategori ini menimbulkan kerugian rata-rata terbanyak, mencapai USD 593 ribu (Aulia, 2024). Beberapa kasus besar semakin mempertegas urgensi isu ini. Misalnya, manipulasi laba oleh PT Garuda Indonesia melalui percepatan pengakuan pendapatan senilai USD 239,94 juta, yang jelas bertentangan dengan prinsip pelaporan keuangan (Christian & Junnestine, 2024). Kasus serupa juga terjadi pada tahun 2023, melibatkan PT Waskita Karya dan PT Wijaya Karya, yang selama

bertahun-tahun menyajikan laporan dengan laba positif meskipun arus kas perusahaan terus negatif (Amelia & Rahman, 2024).

Salah satu model dasar yang digunakan untuk mengidentifikasi indikasi manipulasi adalah Beneish *M-Score* (Beneish, 1999), sebuah model yang memanfaatkan delapan rasio keuangan untuk menilai probabilitas kecurangan laba. Model ini telah menjadi pendekatan populer dalam penelitian mengenai *earnings manipulation* dan terbukti efektif dalam mendeteksi pola pelaporan yang menyimpang (Tarjo & Herawati, 2015). Temuan tersebut selaras dengan penelitian Kurniasari dkk. (2024), yang menunjukkan bahwa Beneish *M-Score* mampu mendeteksi indikasi manipulasi laporan keuangan pada PT Waskita selama periode tahun 2020 hingga 2022.

Meningkatnya kompleksitas transaksi keuangan dan besarnya volume data di era digital membuat metode statistik tradisional semakin kurang memadai untuk menangkap pola-pola nonlinier yang kompleks dalam laporan keuangan. Pendekatan algoritma *Machine Learning* seperti *Logistic Regression*, *Decision Tree*, dan *Random Forest* mampu memberikan akurasi yang lebih tinggi dalam mendeteksi kecurangan. Teknologi ini memungkinkan proses audit dan pemantauan berjalan lebih otomatis, efisien, dan minim subjektivitas, karena model dapat belajar dan presisi (Prasetyo & Dewayanto, 2024).

Efektivitas algoritma *Machine Learning* dalam klasifikasi *fraudulent financial statement* telah banyak dibuktikan. Hajek dan Henriques (2017) menunjukkan bahwa metode berbasis *Decision Tree* memberikan kinerja dengan *Accuracy*, dan sensitivitas lebih tinggi dibandingkan *Logistic Regression*, *Bayesian Model*, *Support Vector Machine* (SVM), dan *Neural Network*. Temuan ini diperkuat

oleh Fu dkk. (2025), yang menyebutkan bahwa *Decision Tree* mampu mencapai akurasi hingga 95 persen dalam klasifikasi *financial fraud*, mengungguli SVM, dan *Logistic Regression*.

Random Forest merupakan salah satu algoritma *Machine Learning* yang banyak digunakan dalam tugas klasifikasi karena mampu menggabungkan sejumlah pohon keputusan dan menghasilkan prediksi yang lebih stabil dan akurat (Breiman, 2001). Keunggulan utama algoritma ini terletak pada kemampuannya mengurangi *overfitting*, mengelola hubungan nonlinier antar variabel, serta bekerja secara efektif pada dataset berukuran besar dengan karakteristik beragam. Namun, pada kasus deteksi kecurangan laporan keuangan atau transaksi, tantangan utama yang sering muncul dari ketidakseimbangan distribusi kelas (*imbalanced data*), jumlah data kategori *fraud* jauh lebih sedikit dibandingkan data kategori *nonfraud*. Kondisi ini dapat menyebabkan model cenderung bias terhadap kelas mayoritas dan mengabaikan pola minoritas yang justru menjadi fokus utama. Untuk mengatasi masalah tersebut, Teknik *Synthetic Minority Oversampling Technique* (SMOTE) umum digunakan dalam meningkatkan representasi kelas minoritas. Penelitian Cheng dkk. (2021) serta Erlin dkk. (2022) menegaskan bahwa algoritma *Random Forest* mampu memberikan performa lebih unggul dalam mendeteksi anomali dan potensi manipulasi laporan keuangan dibandingkan pendekatan linier konvensional, terutama ketika dipadukan dengan teknik penyeimbangan data seperti SMOTE. Temuan-temuan tersebut menunjukkan bahwa integrasi metode *resampling* dengan algoritma *Random Forest* merupakan pendekatan yang efektif untuk meningkatkan akurasi klasifikasi pada konteks *financial fraud detection*, termasuk dalam analisis berbasis rasio keuangan.

Algoritma *Random Forest* yang dikombinasikan dengan SMOTE memang telah terbukti efektif dalam mengatasi ketidakseimbangan data. Namun, terdapat aspek lain yang masih jarang untuk dipertimbangkan dalam proses klasifikasi yaitu penentuan nilai ambang batas (*threshold*) keputusan. Secara *default*, algoritma klasifikasi menggunakan *threshold* sebesar 0,5 untuk mengkonversi probabilitas menjadi label kelas, dimana observasi dengan nilai probabilitas lebih besar dari *threshold* diklasifikasi sebagai kelas positif. Namun, penggunaan *threshold default* ini tidak selalu menghasilkan performa yang optimal pada dataset tidak seimbang (Ehsani & Dastjerdi, 2023).

Threshold optimization bekerja dengan mengevaluasi berbagai nilai *threshold* dalam rentang tertentu dan memilih nilai yang menghasilkan performa terbaik berdasarkan metrik yang ditentukan. Pendekatan ini memiliki keunggulan karena tidak mengubah struktur data asli maupun memodifikasi algoritma pembelajaran, melainkan hanya menyesuaikan keluaran model pada tahap pasca-pemrosesan (Priscilla & Prabha, 2024). Penelitian sebelumnya menunjukkan bahwa *threshold* optimal untuk klasifikasi data tidak seimbang seringkali berbeda secara signifikan dari nilai default 0,5 dan penerapan *threshold optimization* dapat meningkatkan F1-Score secara substansial dibandingkan dengan penggunaan *threshold default* (Zou dkk., 2016).

Meskipun demikian, berbagai kajian sebelumnya masih meninggalkan sejumlah celah penelitian. Sebagian besar studi hanya menguji model *Beneish* dalam pendekatan linier tanpa mengintegrasikannya dengan teknik pembelajaran *Machine Learning* yang mampu menyesuaikan pola data dengan dinamika pasar Indonesia. Selain itu, kajian yang secara khusus menyoroti karakteristik emiten di

Indonesia masih terbatas, padahal pasar berkembang seperti Indonesia memiliki corak pelaporan keuangan yang berbeda dengan pasar global yang cenderung lebih stabil. Oleh karena itu, penelitian ini menghadirkan pembaruan melalui penggabungan Beneish M-Score sebagai variabel input dengan algoritma *Random Forest* yang diperkuat dengan SMOTE untuk mengatasi ketidakseimbangan kelas pada data untuk mendeteksi kecurangan laporan keuangan emiten yang tergabung dalam Indeks Papan Utama Indonesia. Pemilihan emiten yang tergabung pada indeks ini didasarkan pada pertimbangan bahwa perusahaan-perusahaan tersebut umumnya memiliki ukuran besar, rekam jejak yang baik, serta menjadi perhatian utama para investor karena menawarkan potensi keuntungan yang lebih stabil dan prospektif (Rahmawati, 2016).

Urgensi penelitian ini muncul dari kebutuhan akan metode deteksi yang lebih responsif terhadap karakter pasar modal Indonesia. Perilaku investor dalam negeri umumnya masih dipengaruhi keterbatasan informasi publik dan cenderung *herding*, sementara kualitas laporan keuangan sering kali tidak sepenuhnya mencerminkan kondisi ekonomi perusahaan yang sebenarnya. Jika model Beneish konvensional hanya mampu mengidentifikasi potensi manipulasi melalui hubungan liner, maka model *Machine Learning* seperti *Random Forest* mampu menangkap pola nonlinier yang lebih kompleks antar berbagai rasio keuangan dan menghasilkan klasifikasi yang lebih akurat. Melalui penelitian berjudul **“Analisis Kinerja Klasifikasi Algoritma *Random Forest* dilengkapi SMOTE dan *Threshold Optimization* untuk Deteksi Kecurangan Laporan Keuangan Emiten Indeks Papan Utama Indonesia”** ini, diharapkan dapat memberikan alternatif model pendeteksian *fraud* yang dapat mendukung regulator, auditor, dan investor dalam mengidentifikasi

indikasi *fraud* lebih dini, sekaligus memperkuat transparansi serta efisiensi pasar modal nasional.

1.2 Rumusan Masalah

Rumusan masalah berdasarkan latar belakang yang telah dipaparkan sebelumnya, diuraikan sebagai berikut.

1. Bagaimana kinerja algoritma *Random Forest* baik menggunakan *default threshold* maupun *optimal threshold* dalam melakukan klasifikasi laporan keuangan emiten?
2. Apakah penerapan SMOTE pada algoritma *Random Forest* baik menggunakan *default threshold* maupun *optimal threshold* mampu meningkatkan kemampuan dalam melakukan klasifikasi laporan keuangan emiten dibandingkan model dasar algoritma *Random Forest*?

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini dimaksudkan untuk menjaga fokus penelitian dan menghindari perluasan pembahasan yang tidak relevan. Adapun pembatasan dilakukan pada beberapa hal sebagai berikut.

1. Data yang digunakan dalam penelitian merupakan laporan keuangan tahunan publikasi resmi dari emiten nonperbankan yang terdaftar dalam Indeks Papan Utama Bursa Efek Indonesia.
2. Variabel yang digunakan adalah rasio keuangan yang dibentuk berdasarkan pada indikator Beneish M-Score.

3. Algoritma klasifikasi yang digunakan dalam penelitian ini adalah *Random Forest* yang dilengkapi dengan *Synthetic Minority Oversampling Technique* (SMOTE) untuk menangani kelas data yang tidak seimbang.
4. Evaluasi performa model dilakukan menggunakan *Confusion Matrix* dan ukuran *Accuracy*, *Precision*, *Recall*, dan *F1-Score*.

1.4 Tujuan Penelitian

Sejalan dengan rumusan masalah yang telah disampaikan, tujuan dalam penelitian ini sebagai berikut.

1. Membangun model klasifikasi dengan algoritma *Random Forest* untuk digunakan dalam mendeteksi kecurangan dari laporan keuangan emiten
2. Menganalisis pengaruh penerapan SMOTE terhadap peningkatan performa algoritma *Random Forest* dalam proses klasifikasi laporan keuangan emiten.