

Comparative Analysis of the Effectiveness of Three Models in Detecting Financial Statement Fraud

Achmad Haryadi¹

Risna Nurmalasari²

Harti Budi Yanti³

^{1,2,3}Faculty of Economics and Business, Universitas Trisakti, Indonesia

*Correspondences : hartibudi@trisakti.ac.id

ABSTRACT

Financial statement transparency has become a critical issue in the post-Covid-19 period, particularly in the infrastructure, construction, property, and real estate sectors, which face financing pressures and complex revenue recognition. This study aims to compare the effectiveness of the Beneish M-Score, Dechow F-Score, and Altman Z-Score in detecting the risk of financial statement fraud (FSF). A quantitative approach is employed using classification analysis through a confusion matrix, complemented by ROC and AUC testing. The sample consists of 50 companies listed on the Indonesia Stock Exchange during the 2022–2024 period, with a total of 150 financial statement observations. The results indicate that the Altman Z-Score demonstrates the best performance, with an accuracy of 68%, an F1-Score of 62,5%, and an AUC of 0.742, followed by the Dechow F-Score and the Beneish M-Score. These findings suggest that financial distress is a key signal in identifying FSF. The study concludes that the Altman Z-Score is effective as an early warning system for financial statement fraud risk in the examined sectors.

Keywords: Beneish M-Score; Dechow F-Score; Altman Z-Score; Financial Statement Fraud

Analisis Komparatif Efektivitas Tiga Model Dalam Deteksi Financial Statement Fraud

ABSTRAK

Transparansi laporan keuangan menjadi isu krusial pada periode pasca Covid-19, khususnya pada sektor infrastruktur, konstruksi, properti, dan real estate yang menghadapi tekanan pembiayaan dan kompleksitas pengakuan pendapatan. Penelitian ini bertujuan membandingkan efektivitas Beneish M-Score, Dechow F-Score, dan Altman Z-Score dalam mendeteksi risiko financial statement fraud (FSF). Metode penelitian menggunakan pendekatan kuantitatif dengan analisis klasifikasi melalui confusion matrix serta pengujian ROC dan AUC. Sampel terdiri atas 50 perusahaan terdaftar di BEI selama periode 2022–2024 dengan total 150 observasi laporan keuangan. Hasil penelitian menunjukkan Altman Z-Score memiliki kinerja terbaik dengan tingkat akurasi 68%, F1-Score 62,5%, dan AUC 0,742, diikuti oleh Dechow F-Score dan Beneish M-Score. Temuan ini mengindikasikan bahwa tekanan keuangan (financial distress) merupakan sinyal penting dalam mendeteksi FSF. Simpulan penelitian menegaskan Altman Z-Score efektif digunakan sebagai early warning system risiko kecurangan laporan keuangan pada sektor yang diteliti.

Kata Kunci: Beneish M-Score; Dechow F-Score; Altman Z-Score; Financial Statement Fraud

Artikel dapat diakses : <https://ejournal1.unud.ac.id/index.php/Akuntansi/index>



e-ISSN 2302-8556

Vol. 36 No. 2
Denpasar, 28 Februari 2026
Hal. 385-398

DOI:
10.24843/EJA.2026.v36.i02.p09

PENGUTIPAN:

Haryadi, A., Nurmalasari, R.,
& Yanti, H. B. (2026).
Comparative Analysis of the
Effectiveness of Three Models
in Detecting Financial
Statement Fraud.
E-Jurnal Akuntansi,
36(2), 385-398

RIWAYAT ARTIKEL:

Artikel Masuk:
13 November 2025
Artikel Diterima:
6 Januari 2026

PENDAHULUAN

Transparansi dan keandalan laporan keuangan sangat penting bagi pasar modal. Pada periode pasca Covid-19, emiten sektor infrastruktur, konstruksi, properti dan *real estate* di Indonesia menghadapi berbagai tantangan dan tekanan pembiayaan proyek, volatilitas biaya, serta kompleksitas pengakuan pendapatan. Kondisi ini dapat mendorong praktik pelaporan yang agresif dan meningkatkan risiko terjadinya *financial statement fraud*. Dua model deteksi dini atas kecurangan laporan keuangan dengan berbasis data sekunder yang banyak digunakan adalah Beneish M-Score dan Dechow F-Score. Beneish M-Score mengkombinasikan sejumlah rasio keuangan yaitu DSRI, GMI, AQI, SGI, DEPI, SGAI, LVGI, dan TATA (Durana et al., 2022); (Holđa, 2020). Sementara itu, Dechow F-Score mengestimasi risiko kecurangan melalui gabungan indikator keuangan dan non keuangan, dan kerap dibandingkan langsung dengan Beneish di berbagai studi (Ratmono et al., 2020); (Kaab Omeir et al., 2023). Dechow F-Score digunakan untuk memprediksi kemungkinan salah saji (*misstatement*) melalui kombinasi sinyal akrual dan kinerja (Mentari & Indriani, 2024); (Nurcahyono et al., 2021).

Beneish M-score merupakan model statistik yang dikembangkan pada tahun 2019 yang menggunakan rasio keuangan untuk mengevaluasi sejauh mana laba telah dimanipulasi. Beneish menganalisis profil perusahaan yang cenderung melakukan manipulasi laba, kemudian mengembangkan model statistik untuk membedakan antara perusahaan yang melakukan manipulasi dan yang tidak. Selanjutnya, Beneish dan Nichols (2009) menyempurnakan model tersebut dengan menambahkan delapan indikator.

Dechow F-Score adalah model deteksi yang dikembangkan oleh Patricia M. Dechow et al, pada tahun 2011, dan merupakan alat penilaian risiko kecurangan yang menghasilkan skor yang disebut F-Score, yang menjadi indikator probabilitas terjadinya kecurangan pelaporan keuangan (*fraudulent financial reporting*). Dechow dan tim mengembangkan skor prediktif untuk mengidentifikasi perusahaan yang memiliki kemungkinan besar melakukan salah saji material. Dasar pengembangan model ini adalah analisis terhadap seluruh data *Accounting and Auditing Enforcement Releases* (AAER) yang diterbitkan oleh U.S. Securities and Exchange Commission (SEC) selama periode 1982–2005. Nilai F-Score > 1 diartikan risiko salah saji lebih tinggi. Dechow F-Score sebagai ukuran risiko kecurangan yang menggabungkan indikator keuangan dan non-keuangan, nilai F-Score yang lebih tinggi diharapkan berkorelasi positif dengan kemungkinan terdeteksinya kecurangan (Kaab Omeir et al., 2023); (Husnurrosyidah & Fatihah, 2022).

Efektivitas deteksi dini Beneish M-Score dan Dechow F-Score tidak selalu seragam antar negara atau sektor usaha. Penelitian di Malaysia menunjukkan sensitivitas Beneish vs Dechow yang sama-sama kuat yaitu 73,17% vs 69,51%, hal ini menandakan daya deteksi yang berdekatan (Aghghaleh et al., 2016). Hasil penelitian di Afrika Selatan menemukan Beneish vs Dechow sama-sama lemah dalam mendeteksi fraud di Afrika Selatan, sehingga kedua model tersebut tidak cocok diterapkan di Afrika Selatan (Marais et al., 2023). Sedangkan penelitian di Slovakia dengan Beneish 8-Parameter terbukti lebih akurat mengungkap *creative accounting* (Durana et al., 2022).

Penelitian pada sektor perbankan yang terdaftar pada Bursa Efek Indonesia (BEI) periode 2018-2020 dan penelitian pada perusahaan terdaftar di Jakarta Islamic Index (JII) periode 2017-2021 menunjukkan Beneish M-Score lebih efektif untuk deteksi *fraud* laporan keuangan dibandingkan Dechow F-Score (Patmawati & Rahmawati, 2023; Husnurrosyidah & Fatihah, 2022).

Dari perspektif teori agensi, potensi konflik antara pemilik (prinsipal) dan manajer (agen) mendorong timbulnya insentif manipulasi kinerja demi memenuhi target, menjaga reputasi, atau mengamankan kompensasi. Pola insentif ini tercermin pada praktik *earnings management* dan *creative accounting* (Durana et al., 2022).

Dalam konteks Indonesia, pendekatan *Fraud Pentagon* sebagai perluasan *Fraud Triangle* dengan unsur kompetensi dan arogansi menegaskan bagaimana tekanan, peluang, rasionalisasi, kemampuan, dan sikap personal dapat memicu kecurangan pelaporan, kerangka ini kerap dipakai ketika menilai peran F-Score atau M-Score pada data Indonesia (Ratmono et al., 2020).

Sementara itu, teori *signaling* menjelaskan dorongan manajemen untuk mengirim sinyal kinerja yang baik ke pasar melalui angka-angka akuntansi, terutama saat perusahaan menghadapi fase pertumbuhan agresif, *leverage* yang meningkat, atau tekanan proyek jangka panjang. Indikator-indikator inilah yang ditangkap oleh rasio dalam Beneish M-Score, misalnya SGI untuk pertumbuhan penjualan, LVGI untuk perubahan *leverage*, dan TATA sebagai intensitas akrual (Durana et al., 2022).

Penelitian pada perusahaan sektor infrastruktur, konstruksi, properti, dan *real estate* belum banyak dilakukan, khususnya pada periode pasca Covid seiring adanya perubahan pola pengakuan pendapatan kontraktual. Diterbitkannya PSAK 115 tentang Pendapatan dari Kontrak dengan Pelanggan yang berlaku efektif per 1 Januari 2020 akan sangat mempengaruhi bagaimana perusahaan infrastruktur, konstruksi, properti, dan *real estate* dalam mengakui pendapatannya. Hal tersebut sejalan dengan penelitian (Yudhie & Aminah, 2023), dimana penerapan PSAK 115 sebagai standar pengakuan pendapatan memberikan dampak signifikan terhadap kualitas informasi laba, khususnya dalam sektor konstruksi yang memiliki karakteristik proyek jangka panjang dan kompleks.

Berdasarkan kebutuhan tersebut, penelitian ini berupaya menilai kembali efektivitas Beneish M-Score dan Dechow F-Score dalam mendeteksi risiko *financial statement fraud* pada perusahaan sektor infrastruktur, konstruksi, properti, dan *real estate* yang terdaftar sebagai emiten di BEI. Penelitian ini juga menambahkan model ketiga yaitu Altman Z-Score. Meskipun Altman Z-Score dianggap lebih tepat sebagai model deteksi *financial distress*, namun model ini merupakan indikator adanya tekanan keuangan (*pressure*) yang dapat mendorong terjadinya *fraud*. Hasil penelitian (Bhavani & Amponsah, 2017), menunjukkan Altman Z-Score memberikan indikasi bahwa laporan keuangan perusahaan mengandung kelemahan atau penyimpangan. Sementara itu hasil penelitian (Pertiwi, et. Al, 2023) menunjukkan Altman Z-Score memiliki pengaruh yang positif namun tidak signifikan dalam mendeteksi kecurangan keuangan perusahaan BUMN Non Keuangan pada tahun 2016-2020.

METODE PENELITIAN

Tujuan penelitian ini adalah untuk menilai komparasi efektivitas tiga model deteksi yang dapat memprediksi keterjadian ataupun potensi risiko kecurangan laporan keuangan pada periode pasca Covid tahun 2022 s.d. 2024. Perbandingan model deteksi adalah Beneish M-Score, Dechow F-Score, dan Altman Z-Score untuk dinilai mana yang lebih baik dalam mendeteksi kecurangan laporan keuangan.

Objek penelitian ini adalah perusahaan sektor infrastruktur, konstruksi, properti, dan *real estate* yang terdaftar sebagai emiten pada BEI dengan periode pelaporan keuangan tahun 2022 s.d. 2024 (3 tahun). Dari data BEI per Oktober 2025, total perusahaan sektor infrastruktur, konstruksi, properti, dan *real estate* yang tercatat adalah sebanyak 132 perusahaan terdaftar. Untuk penelitian ini, total perusahaan yang diamati adalah sebanyak 50 perusahaan, sehingga total observasi adalah sebanyak 150 laporan keuangan. Pemilihan sampel mengikuti ketersediaan data (*purposive sampling*), yaitu tersedianya laporan keuangan tahunan dalam *annual report* yang di-publish ke publik.

Kecurangan Laporan Keuangan atau *Financial Statement Fraud* (FSF) dalam penelitian ini menggunakan pendekatan observasional, dimana dugaan dan indikasi adanya FSF ditunjukkan secara kuat dengan adanya *restatement* laporan keuangan, *qualified opinion*, sanksi salah saji dan/atau pelaporan laporan keuangan oleh regulator (OJK/BEI), suspensi perdagangan di BEI dengan notasi dan pemantauan khusus, adanya proses Penundaan Kewajiban Pembayaran Utang (PKPU), serta terungkapnya kasus korupsi dan kecurangan pada periode tahun observasi. Perusahaan yang dianggap berisiko melakukan FSF akan diberikan skor 1, sedangkan perusahaan yang dianggap tidak berisiko melakukan manipulasi laporan keuangan akan diberikan skor 0.

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis klasifikasi (*classification analysis*). Metode ini mengevaluasi kemampuan suatu model dalam mengklasifikasikan objek penelitian ke dalam kategori kelas (*fraud* dan *non fraud*). Analisis klasifikasi dilakukan dengan *confusion matrix* yang merupakan alat analisis untuk mengevaluasi performa model klasifikasi, yang diciptakan oleh Karl Pearson pada tahun 1904 (Sathyanarayanan & Tantri, 2024). Matriks ini menyajikan perbandingan antara nilai prediksi model dengan nilai aktual atau observasional berdasarkan kategori fraud (1) dan non fraud (0). Dengan demikian, *confusion matrix* memungkinkan peneliti untuk menilai sejauh mana model dapat mengelompokkan perusahaan secara benar atau salah ke dalam kategori yang ditetapkan. *Confusion matrix* sebagaimana tersaji pada Tabel 1 menjadi dasar untuk menghitung seluruh matrik evaluasi kinerja model.

Tabel 1. Confusion Matrix

Matrix	Actual Fraud (1)	Actual Non-Fraud (0)
Predicted Fraud (1)	TP	FP
Predicted Non-Fraud (0)	FN	TN

Sumber: Sathyanarayanan & Tantri, 2024 (adjusted)

Keterangan:

TP (<i>True Positive</i>)	= aktual fraud, diprediksi fraud
TN (<i>True Negative</i>)	= aktual non-fraud, diprediksi non-fraud
FP (<i>False Positive</i>)	= aktual non-fraud tetapi diprediksi fraud
FN (<i>False Negative</i>)	= aktual fraud tetapi diprediksi non-fraud
<i>Actual Fraud</i>	= fraud yang terjadi berdasarkan hasil observasi
<i>Actual Non Fraud</i>	= non fraud yang terjadi berdasarkan hasil observasi

Matrik evaluasi kinerja selanjutnya diukur dengan beberapa indikator untuk mengetahui performa masing-masing model (Sathyanarayanan & Tantri, 2024), sebagaimana dilihat pada Tabel 2.

Tabel 2. Indikator Matrik Evaluasi Kinerja

Indikator	Rumus
Accuracy	$TP+TN/(TP+TN+FP+FN)$
Sensitivity/Recall	$TP/(TP+FN)$
Specificity	$TN/(TN+FP)$
Precision	$TP/(TP+FP)$
FPR	$FP/(TN+FP)$
F1-Score	$(2 \times \text{Precision} \times \text{Recall})/(\text{Precision} + \text{Recall})$

Sumber: Sathyanarayanan & Tantri, 2024

Dalam penelitian ini, pengujian tambahan juga dilakukan dengan membuat Kurva ROC (*Receiver Operating Characteristic*) dan menghitung luasan AUC (*Area Under the Curve*). Kurva ROC (*Receiver Operating Characteristic*) merupakan visualisasi grafis dari dua parameter, yaitu *True Positive Rate* (TPR) dan *True Negative Rate* (TNR). Kurva ROC dapat membantu dalam memilih ambang batas yang optimal untuk pengklasifikasi guna memaksimalkan positif sejati dan meminimalkan negatif sejati. Plot ideal dari kurva ROC adalah ketika TPR mencapai 100% dengan nol FP. Kurva ROC tepat digunakan untuk mengevaluasi *dataset* yang seimbang (*balanced datasets*). Kurva ROC mampu mengevaluasi kinerja model pada seluruh ambang batas (*thresholds*) klasifikasi, berbeda dengan matrik evaluasi kinerja yang hanya dihitung berdasarkan *cut off* kinerja tertentu. AUC diukur dengan mengukur luas area di bawah kurva ROC dari titik (0,0) sampai (1,1). Nilai AUC berada di antara 0 dan 1. Semakin mendekati 1 menandakan model lebih baik dalam membedakan kelas positif dan negatif. (Sathyanarayanan & Tantri, 2024)

Beneish M-score merupakan model statistik yang menggunakan rasio keuangan untuk mengevaluasi sejauh mana laba telah dimanipulasi. Beneish dan Nichols (2009) menyempurnakan model tersebut dengan menambahkan delapan indikator untuk memperoleh rumus sebagai berikut:

$$\text{M-Score} = -4,84 + (0,92 \times \text{DSRI}) + (0,528 \times \text{GMI}) + (0,404 \times \text{AQI}) + (0,892 \times \text{SGI}) + (0,115 \times \text{DEPI}) - (0,172 \times \text{SGAI}) + (4,679 \times \text{TATA}) - (0,327 \times \text{LVGI}) \dots (1)$$

Dengan indikator dan perhitungan model Beneish M-Score pada Tabel 3, yaitu:
Tabel 3. Indikator Beneish M-Score

Indikator	Rumus
DSRI	$(\text{Receivables } t / \text{Sales } t) / (\text{Receivables } t-1 / \text{Sales } t-1)$
GMI	$\{[\text{Sales } t-1 - (\text{Cost of Good Sold } t-1 / \text{Sales } t-1)] / \text{Sales } t-1\} /$ $[\text{Sales } t - (\text{Cost of Good Sold } t / \text{Sales } t)]$
AQI	$\{1 - [(\text{Curent Assets } t + \text{PPE } t) / \text{Total Assets } t]\} /$ $\{1 - [(\text{Curent Assets } t-1 + \text{PPE } t-1) / \text{Total Assets } t-1]\}$
SGI	$\text{Sales } t / \text{Sales } t-1$
DEPI	$[\text{Depreciation } t-1 / (\text{Depreciation } t-1 + \text{PPE } t-1)] /$ $[\text{Depreciation } t / (\text{Depreciation } t + \text{PPE } t)]$
SGAI	$(\text{Sales, General, Adm. Exp } t / \text{Sales } t) /$ $(\text{Sales, General, Adm. Exp } t-1 / \text{Sales } t-1)$
LVGI	$[\text{Long Term Debt } t + (\text{Current Liabilities } t / \text{Total Asset } t)] /$ $[\text{Long Term Debt } t-1 + (\text{Current Liabilities } t-1 / \text{Total Asset } t-1)]$
TATA	$\text{Total Accruals } t / \text{Total Assets } t$

Sumber: Aghghaleh, et al., 2016

Tahun t merupakan tahun observasi, yaitu 2022, 2023, dan 2024, sementara tahun $t-1$ merupakan tahun sebelum tahun berjalan tersebut. Dari perhitungan rumus tersebut, jika nilai M-Score kurang dari $-2,22$, menunjukkan perusahaan diprediksi tidak melakukan manipulasi laba pada periode tersebut. Sebaliknya, jika nilai M-Score lebih dari $-2,22$, menjadi sinyal perusahaan cenderung melakukan manipulasi laba. (Aghghaleh et al., 2016)

Dechow F-Score adalah model deteksi yang dikembangkan oleh Patricia M. Dechow et al. (2011), dan merupakan alat penilaian risiko kecurangan yang menghasilkan skor yang disebut F-Score, yang menjadi indikator probabilitas terjadinya kecurangan pelaporan keuangan (*fraudulent financial reporting*). Rumus nilai prediksi Dechow F-Score adalah sebagai berikut:

$$\begin{aligned} \text{Predicting Value} = & -7,893 + (0,790 \times \text{RSST}) + (2,518 \times \Delta \text{REC}) + (1,191 \times \Delta \text{INV}) + \\ & (1,979 \times \text{SOFT ASSET}) + (0,171 \times \Delta \text{CASHSALES}) - (0,932 \times \\ & \Delta \text{ROA}) + (1,029 \times \text{ISSUE}) \dots \dots \dots (2) \end{aligned}$$

Indikator dan perhitungan yang digunakan dalam pengukuran Dechow F-Score adalah pada Tabel 4.

Tabel 4. Indikator Dechow F-Score

Indikator	Rumus
RSST	$(\Delta WC + \Delta NCO + \Delta FIN) / \text{Average Total Assets}$ $WC = (\text{Current Assets} - \text{Cash \& Short Term Investment}) - (\text{Current Liabilities} - \text{Debt in Current Liabilities})$ $NCO = (\text{Total Assets} - \text{Current Assets} - \text{Investments \& Advances}) - (\text{Total Liabilities} - \text{Current Liabilities} - \text{Long Term Debt})$ $FIN = (\text{Short Term Investments} + \text{Long Term Investments}) - (\text{Long Term Debt} + \text{Debt in Current Liabilities} + \text{Preferred Stock})$
ΔREC	$\Delta \text{Account Receivables} / \text{Average Total Assets}$
ΔINV	$\Delta \text{Inventory} / \text{Average Total Assets}$
SOFTASSETS	$(\text{Total Assets} - \text{PPE} - \text{Cash \& Cash Equivalents}) / \text{Total Assets}$
$\Delta CASHSALES$	Percentage change in cash sales $(\text{Sales} - \Delta \text{Account Receivables})$
ΔROA	$(\text{Earnings } t / \text{Average total assets } t) - (\text{Earning } t-1 / \text{Average Total Assets } t-1)$
ISSUE	An indicator variable coded 1 if the firm issued securities during year t

Sumber: Aghghaleh, et al., 2016

Atas nilai prediksi tersebut belum menunjukkan probabilitas terjadinya manipulasi, sehingga nilai *predicting value* dikonversi menjadi probabilitas dengan rumus:

$$\text{Probability} = \frac{e^{\text{Predicted Value}}}{1 + e^{\text{Predicted Value}}} \dots \dots \dots (3)$$

Selanjutnya dibagi dengan probabilitas tanpa syarat dari salah saji (*misstatement*) sebesar 0,0037 untuk memperoleh F-Score.

$$\text{F Score} = \frac{\text{Probability}}{0,0037} \dots \dots \dots (4)$$

F-Score yang lebih besar dari 1 menunjukkan probabilitas salah saji yang lebih tinggi dibandingkan probabilitas tanpa syarat tersebut. (Aghghaleh, et al., 2016). F-Score > 1 menunjukkan probabilitas kemungkinan perusahaan melakukan manipulasi laporan keuangan yang tinggi. Semakin tinggi F-Score, semakin besar kemungkinan laporan keuangan mengandung salah saji material atau fraud.

Altman Z-Score adalah model yang dikembangkan oleh Edward I. Altman (1968) untuk memprediksi risiko kebangkrutan (*financial distress*) pada perusahaan manufaktur publik. Meskipun awalnya diciptakan untuk menilai kesehatan keuangan dan memprediksi risiko kebangkrutan, Altman Z-Score juga dapat memberikan indikasi awal adanya manipulasi laporan keuangan sebagai proxy unsur tekanan/*pressure*. Rumus Altman Z-Score dalam penelitian ini menggunakan rumus modifikasi Z'' yang dianggap cocok digunakan untuk berbagai tipe industri. (Diep, T. T., & Phung, V. T. H, 2019).

$$Z'' = 6,56X1 + 3,26X2 + 6,72X3 + 1,05X4 \dots \dots \dots (5)$$

Indikator dan perhitungan yang digunakan dalam pengukuran Altman Z-Score modifikasi sebagaimana Tabel 5.

Tabel 5. Indikator Altman Z-Score

Indikator	Rumus
X ₁	Working Capital/Total Assets
X ₂	Retained Earnings/Total Assets
X ₃	EBIT/Total Assets
X ₄	Market Value of Equity/Book Value of Total Liabilities

Sumber: Diep, T. T., & Phung, V. T. H, 2019

Hasil pengukuran Z-Score >2,6 menunjukkan perusahaan masih berada di zona aman atau termasuk perusahaan sehat. Nilai Z-Score antara 1,1 sampai dengan 2,6 menunjukkan perusahaan berada pada zona abu-abu dan berpotensi distress. Sementara itu, nilai Z-Score <1,1 menunjukkan perusahaan berada dalam zona bahaya dan risiko *financial distress* atau kebangkrutannya tinggi. (Diep, T. T., & Phung, V. T. H, 2019).

Tabulasi hasil perhitungan skor dari masing-masing model dilakukan dalam suatu kertas kerja perhitungan, untuk selanjutnya dilakukan langkah analisis klasifikasi. Jika diketahui terdapat perusahaan yang melakukan *restatement* laporan keuangan di tahun observasi, maka perhitungan menggunakan angka sebelum dilakukan penyajian kembali tersebut. Hal ini dilakukan untuk dapat menangkap sinyal manipulasi dan salah saji material.

Perusahaan sektor infrastruktur, konstruksi, properti, dan *real estate* yang diamati dalam 3 tahun adalah sebanyak 50 perusahaan terdiri dari 11 BUMN dan 39 perusahaan swasta, dengan karakteristik *firm size* dari segi nilai aset dan *sales* yang digambarkan secara statistik deskriptif pada Tabel 6.

Tabel 6. Statistik Deskriptif Firm Size

Descriptive Statistics Firm Size						
	N	Range	Minimum	Maximum	Mean	Std. Deviation
Assets (million Rp)	150	140.695.837	3.0602	140.726.439	17.458.236.633	25.606.852.589
Sales (million Rp)	150	28.700.485	2.727	28.703.212	3.929.058.307	5.852.696.442
Valid N (listwise)	150					

Sumber: Data Penelitian, 2025

Nilai total aset dari 150 sampel menunjukkan *range* sebesar Rp140.695.837 (juta), dengan nilai minimum Rp30.602 (juta) dan nilai maksimum Rp140.726.439 (juta). Rata-rata (*mean*) total aset perusahaan adalah Rp17.458.236,63 (juta) dengan standar deviasi sebesar Rp25.606.852,59 (juta). Standar deviasi yang lebih besar dibandingkan nilai rata-rata mengindikasikan adanya variasi ukuran perusahaan yang cukup tinggi dalam sampel penelitian. Sementara itu, nilai *sales* memiliki rentang sebesar Rp28.700.485 (juta), dengan nilai minimum Rp2.727 (juta) dan nilai

maksimum Rp28.703.212 (juta). Nilai rata-rata penjualan perusahaan adalah Rp3.929.058,31 (juta) dengan standar deviasi sebesar Rp5.852.696,44 (juta), yang juga menunjukkan tingkat variasi penjualan yang relatif tinggi antar perusahaan. Sampel penelitian telah mencerminkan perusahaan dengan ukuran yang heterogen, baik dilihat dari sisi total aset maupun penjualan, sehingga diharapkan mampu merepresentasikan kondisi perusahaan dengan skala kecil hingga besar.

HASIL DAN PEMBAHASAN

Hasil tabulasi data atas skor Beneish, Dechow, dan Altman atas 150 sampel laporan keuangan dapat digambarkan dengan data statistik deskriptif pada Tabel 7.

Tabel 7. Statistik Deskriptif Skor Masing-Masing Model

	Descriptive Statistics						Std.
	N	Range	Minimum	Maximum	Mean	Deviation	
Beneish M-Score	150	96,329	-5,992	90,337	-1,526	7,689	59,129
Dechow F-Score	150	11,476	,0137	11,490	1,060	1,547	2,396
Altman Z-Score	150	65,509	-11,394	54,115	4,585	8,815	77,722
Valid N (listwise)	150						

Sumber: Data Penelitian, 2025

Nilai Beneish M-Score memiliki rentang sebesar 96,329, dengan nilai minimum -5,992 dan maksimum 90,337. Rata-rata M-Score tercatat sebesar -1,526, dengan standar deviasi 7,689, yang menunjukkan adanya perbedaan skor yang sangat kontras antar perusahaan dalam penelitian. Standar deviasi yang jauh lebih besar dari rata-rata menunjukkan bahwa data cenderung tersebar luas dan tidak mengumpul pada satu titik nilai. Dikarenakan perhitungan skor model ini menggunakan data laporan keuangan yang telah *publish*, maka atas *outlier* Beneish yang diketahui terjadi pada satu sampel laporan keuangan tidak dilakukan eliminasi data. Penelitian ini mempertahankan seluruh observasi termasuk nilai ekstrem (*outlier*) untuk mencerminkan kondisi empiris yang sesungguhnya.

Nilai Dechow F-Score menunjukkan rentang sebesar 11,476, dengan nilai minimum 0,0137 dan maksimum 11,490. Rata-rata F-Score adalah 1,060, dengan standar deviasi 1,547. Dibandingkan model lainnya, Dechow F-Score memiliki variasi data yang paling sempit dengan *range* hanya 11,476 dan varians 2,396. Karena standar deviasinya relatif kecil, data pada model Dechow cenderung lebih konsisten atau memiliki sebaran yang lebih rapat dibandingkan dua model lainnya.

Nilai Altman Z-Score memiliki rentang sebesar 65,509, dengan nilai minimum -11,394 dan maksimum 54,115. Rata-rata Z-Score tercatat sebesar 4,585, dengan standar deviasi 8,815. Standar deviasi yang besar mencerminkan heterogenitas kondisi keuangan perusahaan dalam sampel penelitian. Altman Z-Score memiliki tingkat *variance* tertinggi (77,722), yang mencerminkan perbedaan yang lebar atas tingkat kesehatan finansial antar perusahaan yang diteliti.

FSF observasional tahun 2022-2024 dalam penelitian ini adalah sebanyak 57 laporan keuangan yang diindikasikan berisiko fraud, dan 93 tidak berisiko fraud sebagaimana tersaji pada gambar 1. FSF observasional ini akan menjadi *actual fraud* dan *actual non-fraud*.

Case Processing Summary	
FSF	Valid N (listwise)
Positive ^a	57
Negative	93
Missing	0
Total	150

Larger values of the test result variable(s) indicate stronger evidence for a positive actual state.

a. The positive actual state is 1.

Gambar 1. Hasil FSF Observasional

Sumber: Data Penelitian, 2025

Perhitungan *confusion matrix* dilakukan dengan *cut-off* kriteria dari masing-masing model. Dari hasil perhitungan skor dan kesesuaian dengan *cut-off* kriteria, perusahaan yang diprediksi fraud, akan diberikan skor 1, dan perusahaan yang diprediksi sebagai non fraud akan diberikan skor 0. Selanjutnya hasil prediksi tersebut disajikan dalam *confusion matrix* dengan menyandingkan dengan hasil observasi data FSF. *Confusion matrix* ini menyajikan data *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) atau *Type I error*, dan *False Negative* (FN) atau *Type II error* sebagaimana tersaji pada Tabel 8.

Tabel 8. Hasil Confusion Matrix 3 Model

Matrix	Beneish M-Score			Dechow F-Score			Altman Z-Score		
	Actual Fraud	Actual Non-Fraud	Total	Actual Fraud	Actual Non-Fraud	Total	Actual Fraud	Actual Non-Fraud	Total
Pred. Fraud	23	28	57	17	17	57	40	31	57
Pred. Non-Fraud	34	65	93	40	76	93	17	62	93
Total	57	93	150	57	93	150	57	93	150

Sumber: Data Penelitian, 2025

Dari data *confusion matrix*, selanjutnya dilakukan perhitungan pengukuran matrik evaluasi kinerja dengan hasil pada Tabel 9.

Tabel 9. Hasil Perhitungan Matrik Evaluasi Kinerja

Matrik	Beneish M-Score	Dechow F-Score	Altman Z-Score
Accuracy	58,67%	62,00%	68,00%
Sensitivity/Recall	40,35%	29,82%	70,18%
Specificity	69,89%	81,72%	66,67%
Precision	45,10%	50,00%	56,34%
FPR (False Positive Rate)	30,11%	18,28%	33,33%
F1-Score	42,59%	37,36%	62,50%

Sumber: Data Penelitian, 2025

Hasil evaluasi kinerja menggunakan *confusion matrix* menunjukkan adanya *trade-off* yang jelas antara kemampuan mendeteksi fraud (*sensitivity*) dan kemampuan mengidentifikasi non-fraud (*specificity*) pada ketiga model yang diuji.

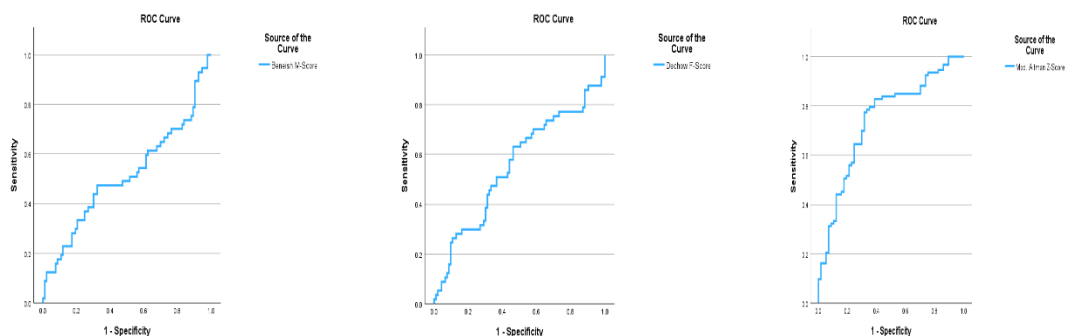
Beneish M-Score menunjukkan kinerja yang relatif moderat, namun paling lemah dibandingkan dua model lainnya. Dengan *accuracy* sebesar 58,67% dan

sensitivity sebesar 40,35%, model ini hanya mampu mendeteksi sebagian kecil kasus fraud. Nilai *precision* sebesar 45,10% menunjukkan bahwa hampir setengah dari prediksi fraud tidak akurat. Selain itu, FPR sebesar 30,11% menandakan tingkat kesalahan yang masih cukup tinggi. Nilai F1-Score sebesar 42,59% menunjukkan bahwa kinerja Beneish M-Score berada di bawah Altman Z-Score dan tidak seimbang antara deteksi fraud dan non-fraud.

Dechow F-Score menunjukkan kinerja yang lebih konservatif. Model ini memiliki *specificity* tertinggi sebesar 81,72%, yang berarti sangat baik dalam mengidentifikasi perusahaan non-fraud. Selain itu, FPR terendah sebesar 18,28% menunjukkan tingkat kesalahan klasifikasi non-fraud yang paling kecil. Namun, kelemahan utama Dechow F-Score terletak pada *sensitivity* yang rendah (29,82%), sehingga model ini kurang efektif dalam menangkap sinyal FSF. Nilai F1-Score sebesar 37,36% menegaskan bahwa model ini lebih cocok digunakan untuk meminimalkan *false alarm* deteksi fraud.

Altman Z-Score menunjukkan kinerja terbaik secara keseluruhan. Model ini memiliki *accuracy* tertinggi sebesar 68,00%, yang mengindikasikan kemampuan klasifikasi yang relatif lebih baik dibandingkan dua model lainnya. Selain itu, *sensitivity* Altman Z-Score mencapai 70,18%, yang berarti model ini paling efektif dalam mengidentifikasi perusahaan yang melakukan FSF. Nilai F1-Score sebesar 62,50% juga menunjukkan keseimbangan yang baik antara kemampuan mendeteksi fraud dan ketepatan prediksi, sehingga model ini paling andal untuk tujuan *early warning* FSF. Namun demikian, nilai *false positive rate* (FPR) sebesar 33,33% mengindikasikan bahwa masih terdapat risiko salah klasifikasi perusahaan non-fraud sebagai fraud.

Dari dataset skor tiga model yang sudah ditabulasi, selanjutnya dilakukan pengujian tambahan berupa grafis ROC dan luasan AUC. Sebagaimana diketahui bahwa untuk model Beneish M-Score dan Dechow F-Score, semakin tinggi skor, maka semakin berisiko fraud, sehingga *state value* ROC Beneish dan Dechow adalah 1. Sebaliknya untuk model Altman Z-Score, semakin tinggi skor, maka semakin rendah untuk terjadinya risiko fraud, sehingga *state value* ROC Altman adalah 0. ROC dari tiga model tersebut tersaji pada Gambar 2.



ROC Beneish M-Score

ROC Dechow F-Score
Gambar 2. Model Score

ROC Altman Z-Score

Sumber: Data Penelitian, 2025

Tabel 10. Hasil Perhitungan Luasan AUC

Area Under the ROC Curve (AUC)		
Test Result Variable(s): Beneish M-Score	Test Result Variable(s): Dechow F-Score	Test Result Variable(s): Altman Z-Score
Area	Area	Area
0.513	0.554	0.742

Sumber: Data Penelitian, 2025

Nilai luasan AUC merepresentasikan kemampuan masing-masing model dalam membedakan perusahaan bermasalah dan tidak bermasalah pada seluruh rentang *threshold* klasifikasi. Hasil perhitungan AUC menunjukkan bahwa Altman Z-Score (0,742) memiliki nilai AUC tertinggi, diikuti Dechow F-Score (0,554) dan Beneish M-Score (0,513). Nilai AUC Altman Z-Score menunjukkan model ini memiliki kemampuan klasifikasi yang lebih baik dan efektif dibandingkan Beneish M-Score maupun Dechow F-Score.

SIMPULAN

Hasil penelitian dengan pengujian komparatif terhadap tiga model deteksi FSF pada perusahaan sektor infrastruktur, konstruksi, properti, dan *real estate* di Indonesia periode 2022-2024, baik dengan metode matrik evaluasi kinerja, maupun ROC dan AUC, menunjukkan hasil yang konsisten bahwa model paling efektif untuk menangkap sinyal FSF adalah Altman Z-Score dengan tingkat akurasi 68%, F1-Score 62,5%, dan AUC 0,742. Selanjutnya diikuti dengan Dechow F-Score dan Beneish M-Score. Meskipun Altman Z-Score merupakan model deteksi *financial distress*, namun model ini ternyata juga dapat digunakan sebagai alat deteksi atau *early warning* keterjadian FSF. Hal tersebut sejalan dengan *fraud theory* bahwa tekanan (*distress*) keuangan dapat memicu tindakan FSF. Penelitian ini memiliki beberapa keterbatasan antara lain penentuan fraud aktual yang bersumber dari justifikasi hasil observasi *restatement*, *red flags*, *sanction-based detection* oleh regulator, dan informasi publik lainnya. Pendekatan ini dapat memiliki *element of subjectivity*, namun hal ini dapat dihindari dengan justifikasi penelitian yang objektif. Selain itu, dengan sampel penelitian berupa perusahaan sektor infrastruktur, konstruksi, properti, dan *real estate* di Indonesia, dimungkinkan memiliki beberapa karakteristik akun laporan keuangan yang unik sehingga membuat kinerja atau performa model deteksi FSF, seperti Beneish dan Dechow menjadi kurang maksimal jika diterapkan pada sektor usaha ini. Oleh karena itu, generalisasi hasil dari penelitian ini harus dilakukan dengan hati-hati.

REFERENSI

Adhyapak Sahayak, H L College of Commerce, Gujarat University, Ahmedabad: 380009, & N. Parikh, Dr. P. (2024). Evaluating Financial Stability and Fraud Detection in Selected Indian Pharmaceutical Companies Using Altman Z Score and Beneish M Score Models. *International Journal of Social Science Humanity & Management Research*, 3(06).
<https://doi.org/10.58806/ijsshmr.2024.v3i6n30>

- Adoboe-Mensah, N., Salia, H., & Addo, E. B. (2023). Using the Beneish M-score Model to Detect Financial Statement Fraud in the Microfinance Industry in Ghana. *International Journal of Economics and Financial Issues*, 13(4), <https://doi.org/10.32479/ijefi.14489>
- Aghghaleh, S. F., Mohamed, Z. M., & Rahmat, M. M. (2016). Detecting Financial Statement Frauds in Malaysia: Comparing the Abilities of Beneish and Dechow Models. *Asian Journal of Accounting and Governance*, 7, 57–65. <https://doi.org/10.17576/AJAG-2016-07-05>
- Anggraeni, F. N. & Tarjo. (2024). Analisis Beneish M-Score Dalam Mendeteksi Indikasi Fraud Laporan Keuangan. *APSSAI Accounting Review*, 4(1), 46–64. <https://doi.org/10.26418/apssai.v4i1.104>
- Bhavani, G & Amponsah, C.T. (2017). M-Score and Z-Score For Detection of Accounting Fraud. *Accountancy Business and the Public Interest*, 1(1), 68-86.
- Beneish, M. D. (1999). The detection of earnings manipulation. *Financial Analysts Journal*, 55(5), 24–36. <https://doi.org/10.2469/faj.v55.n5.2296>.
- Dechow, P. M., Ge, W., Larson, C. R., & Sloan, R. G. (2011). *Predicting material accounting misstatements*. *Contemporary Accounting Research*, 28(1), 17–82. <https://doi.org/10.1111/j.1911-3846.2010.01041.x>.
- Diep, T. T., & Phung, V. T. H. (2019). An application of Altman Z-score model to analyze the bankruptcy risk: Cases of multidisciplinary enterprises in Vietnam. *Investment Management and Financial Innovations*, 16(4), 181–191. [https://doi.org/10.21511/imfi.16\(4\).2019.16](https://doi.org/10.21511/imfi.16(4).2019.16).
- Durana, P., Blazek, R., Machova, V., & Krasnan, M. (2022). The use of Beneish M-scores to reveal creative accounting: Evidence from Slovakia. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 17(2), 481–510. <https://doi.org/10.24136/eq.2022.017>
- Hołda, A. (2020). Using the Beneish M-score model: Evidence from non-financial companies listed on the Warsaw Stock Exchange. *Investment Management and Financial Innovations*, 17(4), 389–401. [https://doi.org/10.21511/imfi.17\(4\).2020.33](https://doi.org/10.21511/imfi.17(4).2020.33)
- Hugo, J. (2019). Efektivitas Model Beneish M-Score Dan Model F-Score Dalam Mendeteksi Kecurangan Laporan Keuangan. *Jurnal Muara Ilmu Ekonomi dan Bisnis*, 3(1), 165. <https://doi.org/10.24912/jmieb.v3i1.2296>
- Husnurrosyidah, H., & Fatihah, I. (2022). Fraud Detecting Using Beneish M-Score and F-Score: Which is More Effective? *Equilibrium: Jurnal Ekonomi Syariah*, 10(1), 137. <https://doi.org/10.21043/equilibrium.v10i1.15351>
- Kaab Omeir, A., Vasiliauskaite, D., & Soleimanizadeh, E. (2023). Detection of financial statements fraud using Beneish and Dechow models. *Journal of Governance and Regulation*, 12(3, special issue), 334–344. <https://doi.org/10.22495/jgrv12i3siart15>
- Magdalena, F., & Tanusdjaja, H. (2018). Analisis Komparasi Metode Altman Z-Score – Financial Ratio dan Metode Beneish M-Score Model – Data Mining dalam Mendeteksi Fraudulent Financial Reporting. *Jurnal Muara Ilmu Ekonomi dan Bisnis*, 2(1), 14. <https://doi.org/10.24912/jmieb.v2i1.1530>

- Marais, A., Vermaak, C., & Shewell, P. (2023). Predicting financial statement manipulation in South Africa: A comparison of the Beneish and Dechow models. *Cogent Economics & Finance*, 11(1), 2190215. <https://doi.org/10.1080/23322039.2023.2190215>
- Marbun, E. (2022). Analisis Prediksi Kebangkrutan Dengan Altman Z-Score Dan Deteksi Kecurangan Laporan Keuangan Dengan Beneish M-Score Pada Perusahaan Asuransi. *Jurnal Cafetaria*, 3(2), 135–150. <https://doi.org/10.51742/akuntansi.v3i2.64>
- Mentari, C., & Indriani, E. (2024). Deteksi Fraudulent Financial Statement melalui Dechow F-Score dengan Pemoderasi Firm Size. *Owner*, 8(4), 4436–4448. <https://doi.org/10.33395/owner.v8i4.2361>
- Miharsi, D., Gamayuni, R. R., & Dharma, F. (2023). Analysis Of The Utilization Of Altman Z-Score, Beneish M-Score, And F-Score Model In Detecting Fraudulent Of Financial Reporting: A Literature Review. *Marginal Journal Of Management Accounting General Finance And International Economic Issues*, 3(2), 353–364. <https://doi.org/10.55047/marginal.v3i2.954>
- Nurchayono, N., Hanum, A. N., Kristiana, I., & Pamungkas, I. D. (2021). Predicting Fraudulent Financial Statement Risk: The Testing Dechow F-Score Financial Sector Company Indonesia. *Universal Journal of Accounting and Finance*, 9(6), 1487–1494. <https://doi.org/10.13189/ujaf.2021.090625>
- Patmawati, P., & Rahmawati, M. (2023). Deteksi Financial Statement Fraud: Model Beneish M-Score, dan Model F-Score. *E-Jurnal Akuntansi*, 33(1), 34. <https://doi.org/10.24843/EJA.2023.v33.i01.p03>
- Pertiwi, J. C., Oktavia, R., & Amelia, Y. (2023). Analisis perbandingan metode pendeteksian kecurangan keuangan menggunakan Altman Z-Score, Beneish M-Score, dan Springate. *Fair Value: Jurnal Ilmiah Akuntansi dan Keuangan*, 5(6), 2666–2676. <https://doi.org/10.32670/fairvalue.v5i6.2534>
- Puspitha, M. Y., & Diantini, N. N. A. (2023). Kemampuan Beneish M-Score Dalam Memprediksi Fraudulent Financial Reporting. *E-Jurnal Akuntansi*, 33(2), 570. <https://doi.org/10.24843/EJA.2023.v33.i02.p20>
- Ratmono, D., Darsono, D., & Cahyonowati, N. (2020). Financial Statement Fraud Detection With Beneish M-Score and Dechow F-Score Model: An Empirical Analysis of Fraud Pentagon Theory in Indonesia. *International Journal of Financial Research*, 11(6), 154. <https://doi.org/10.5430/ijfr.v11n6p154>
- Sathyanarayanan, S., & Tantri, B. R. (2024). Confusion Matrix-Based Performance Evaluation Metrics. *African Journal of Biomedical Research*, 27(4S), 4023–4031. <https://doi.org/10.53555/AJBR.v27i4S.4345>
- Valaskova, K., & Fedorko, R. (2021). Beneish M-score: A measure of fraudulent financial transactions in global environment? *SHS Web of Conferences*, 92, 02064. <https://doi.org/10.1051/shsconf/20219202064>
- Yudhie, N. F. P., & Aminah, I. (2025). Analisis dampak penerapan PSAK 115 terhadap kualitas laba perusahaan subsektor konstruksi di Bursa Efek Indonesia tahun 2020–2024.