

PREDIKSI RISIKO CRASH HARGA SAHAM BERDASARKAN KETERBUKAAN DATA PUBLIK MENGGUNAKAN PENDEKATAN MACHINE LEARNING

Aji Susanto¹, Martanto², Denni Pratama³, Mulyawan⁴.

Program Studi Teknik Informatika¹
Program Studi Manajemen Informatika²
Program Studi Komputerisasi Akuntansi³
Program Studi Sistem Informasi⁴

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
barraghava1@gmail.com

(*) Corresponding Author : barraghava1@gmail.com
Published : 20 Juli 2026

Abstract—This study develops and evaluates a predictive framework for identifying the risk of sharp stock price declines using machine learning techniques and publicly available data. An LSTM-based Deep Neural Network model and several conventional models (Logistic Regression, Random Forest, XGBoost) were used for comparison in the tested approach. Daily market indicators, macro variables, and public disclosure indicators are all components included in the dataset used. Data preprocessing (cleaning, imputation, normalization, time-series framing), feature engineering, hyperparameter optimization (Bayesian/Optuna for DNN), training with rolling-window validation, and interpretability analysis using XAI techniques (SHAP/LIME). Experimental results reveal the “accuracy paradox”: although all models exhibit high accuracy on the test data, extreme class recall imbalance and near-zero F1 scores for the accident class render the models incapable of detecting accident events (the minority class). Under imbalanced data conditions without special treatment, complex LSTM architectures do not demonstrate significant advantages over simple models. The Disclosure Index variable contributes marginally, while feature contribution analysis indicates the dominance of technical signals, such as daily returns. The results show that the use of evaluation metrics emphasizing sensitivity to extreme events (recall, F1, PR-AUC), handling data imbalance (such as oversampling/SMOTE, cost-sensitive learning), and enriching open-ended variables with qualitative signals (NLP/sentiment) is crucial. This study suggests that investors and regulators should exercise caution when using ML models without verifying their sensitivity to extreme events. They should also encourage further research that enhances the capabilities of early warning systems by incorporating alternative data and class balancing strategies. The results indicate that the use of evaluation metrics emphasizing sensitivity to extreme events (recall, F1, PR-AUC), handling class imbalance (such as oversampling/SMOTE, cost-sensitive learning, and enriching feature sets with qualitative signals (NLP/sentiment)), is crucial. This study suggests that investors and regulators should exercise caution when using ML models without verifying their sensitivity to extreme outcomes. They should also encourage further research to improve the capabilities of early warning systems by incorporating alternative data and class balancing strategies.

Keywords: stock price crash risk; public disclosure; deep learning (LSTM); class imbalance; model interpretability.

Abstrak—Penelitian ini mengembangkan dan mengevaluasi kerangka prediktif untuk mengidentifikasi risiko penurunan harga saham tajam dengan menggunakan teknik machine learning dan data publik terbuka. Model Deep Neural Network berbasis LSTM dan beberapa model konvensional (Logistic Regression, Random Forest, XGBoost) digunakan sebagai perbandingan dalam pendekatan yang diuji. Indikator pasar frekuensi harian, variabel makro, dan indikator keterbukaan publik/pengungkapan publik adalah semua komponen yang digunakan dalam dataset yang digunakan. Pra-pemrosesan data (pembersihan, imputasi, normalisasi, framing deret waktu), rekayasa fitur, optimasi hiperparameter (bayesian/optuna untuk DNN), pelatihan dengan validasi rolling-window, dan analisis interpretabilitas

menggunakan teknik XAI (SHAP/LIME). Hasil eksperimen menunjukkan fenomena "akurasi paradoks": meskipun semua model memiliki akurasi tinggi pada data uji, ketidakseimbangan recall kelas yang ekstrim dan skor F1 yang hampir nol untuk kelas kecelakaan membuat model tidak mampu menemukan kejadian kecelakaan (kelas minoritas). Dalam kondisi data yang timpang tanpa perlakuan khusus, arsitektur LSTM yang kompleks tidak menunjukkan keuntungan yang signifikan dibandingkan dengan model sederhana. Variabel Disclosure Index memberikan kontribusi marginal, sementara analisis kontribusi fitur menunjukkan dominasi sinyal teknikal, seperti return harian. Hasil menunjukkan bahwa penggunaan metrik evaluasi yang menekankan sensitivitas terhadap kejadian ekstrim (recall, F1, PR-AUC), penanganan imbalance (seperti oversampling/SMOTE, pembelajaran yang sensitif terhadap biaya, dan pengayaan variabel keterbukaan dengan sinyal kualitatif (NLP/sentimen) sangat penting. Studi ini menyarankan agar investor dan regulator berhati-hati saat menggunakan model ML tanpa memverifikasi sensitivitas terhadap peristiwa ekstrim. Mereka juga harus mendorong penelitian lanjutan yang meningkatkan kemampuan sistem peringatan dini dengan menggabungkan data alternatif dan strategi penyeimbangan kelas.

Kata Kunci: risiko crash saham; keterbukaan data publik; deep learning (LSTM); ketidakseimbangan kelas; interpretabilitas model.

INTRODUCTION

Risiko penurunan harga saham (stock price crash risk) telah menjadi isu penting dalam literatur keuangan modern karena dampaknya terhadap stabilitas pasar modal dan kepercayaan investor. Fenomena ini menggambarkan kondisi ketika harga saham suatu perusahaan mengalami penurunan tajam secara tiba-tiba akibat akumulasi informasi negatif yang tertunda atau tersembunyi oleh manajemen (bad news hoarding) (Shi et al., 2023). Transparansi informasi menjadi kunci untuk meminimalkan risiko tersebut. Ketika perusahaan gagal mengungkapkan informasi secara tepat waktu dan akurat, pasar bereaksi negatif, sehingga meningkatkan volatilitas dan potensi crash (Ge et al., 2023).

Dalam konteks pasar modal China, sejumlah penelitian empiris telah menunjukkan bahwa praktik pengungkapan informasi bisnis memiliki dampak yang signifikan terhadap kemungkinan harga saham jatuh. Tata kelola perusahaan dan pengungkapan informasi lingkungan dapat meningkatkan kepercayaan investor dan mengurangi asimetri informasi (Xu et al., 2022; Ge et al., 2023). Selain itu, transparansi rantai pasokan dan pelaporan risiko dalam diskusi dan analisis manajemen (MD&A) telah terbukti membantu mengurangi kemungkinan kejatuhan harga saham (Zhong et al., 2024; Su et al., 2023). Sebaliknya, pengungkapan informasi yang "malas" atau tidak lengkap meningkatkan risiko crash karena ketidakpastian di pasar meningkat (Shi et al., 2023).

Sebaliknya, metode baru untuk mengidentifikasi dan memprediksi risiko penurunan harga saham dihasilkan dari pengembangan kecerdasan buatan (AI) dan

pembelajaran mesin. Metode ini dapat menemukan pola non-linear antara atribut bisnis dan risiko kecelakaan dengan menggunakan analisis data besar dan variabel keuangan (Li et al., 2024). Misalnya, telah terbukti bahwa prediksi risiko crash di pasar modal China lebih akurat dengan menggunakan indikator keuangan dan data perilaku manajemen (Deng et al., 2022).

Selain itu, penelitian yang dilakukan oleh Jiang et al. (2024) menemukan bahwa algoritma pembelajaran mesin seperti Random Forest atau Gradient Boosting dapat digunakan secara efektif untuk menganalisis atribut penting perusahaan dan elemen manajemen. Algoritma ini dapat digunakan untuk memprediksi kemungkinan penurunan harga saham. Hasil penelitian Karasan et al. (2025) memperkuat temuan ini dengan menyarankan sistem pembelajaran mesin berbasis analisis risiko operasional yang dapat digunakan di berbagai sektor. Bahkan, penelitian terbaru menggunakan analisis teks laporan keuangan dan berita pasar untuk memperkirakan risiko penurunan harga saham perbankan (Prediksi risiko penurunan harga saham bank, 2025). Oleh karena itu, ada peluang penelitian yang signifikan ketika teknologi machine learning dan keterbukaan informasi perusahaan digabungkan. Metode ini membantu menciptakan model prediksi keuangan yang lebih akurat dan memperkuat sistem tata kelola perusahaan melalui peningkatan akuntabilitas dan transparansi (Li et al., 2024; Karasan et al., 2025).

Penelitian ini diharapkan dapat menjelaskan hubungan antara pengungkapan informasi, fitur manajemen, dan risiko crash harga saham dengan memanfaatkan pendekatan data-driven dan teori keuangan perilaku. Fokus ini penting untuk meningkatkan pemahaman kita tentang mengurangi risiko pasar, terutama dalam konteks ekonomi digital yang dipenuhi dengan ketidakpastian dan kompleksitas data.

MATERIALS AND METHODS

Diagram alir penelitian berikut menggambarkan tahapan sistematis dalam penelitian prediksi risiko crash harga saham berdasarkan keterbukaan data publik menggunakan pendekatan machine learning. Alur penelitian disusun secara terstruktur mulai dari identifikasi masalah hingga penyusunan kesimpulan dan saran penelitian. Penelitian ini dilatarbelakangi oleh pentingnya kemampuan dalam mendeteksi potensi penurunan harga saham secara ekstrem agar dapat membantu investor, regulator, maupun perusahaan dalam pengambilan keputusan yang lebih tepat. Dengan memanfaatkan data publik, data keuangan perusahaan, serta teknik deep learning dan machine learning, penelitian ini berupaya menghasilkan model prediksi yang akurat, robust, dan mampu memberikan interpretasi terhadap faktor-faktor yang mempengaruhi risiko crash saham.



Gambar 1. Alur penelitian

Gambar diagram alir penelitian tersebut menunjukkan proses penelitian yang terdiri atas delapan tahapan utama yang saling terhubung secara sistematis. Tahap pertama dimulai dengan identifikasi dan perumusan masalah yang mencakup latar belakang penelitian, rumusan masalah, tujuan penelitian, manfaat

penelitian, serta batasan penelitian. Pada tahap ini ditentukan fokus penelitian terkait prediksi risiko crash harga saham menggunakan pendekatan machine learning berdasarkan keterbukaan data publik. Tahap kedua adalah studi literatur dan kerangka konseptual yang berisi tinjauan pustaka mengenai risiko crash saham, keterbukaan data publik, machine learning, serta penelitian terdahulu yang relevan. Dari tahapan ini disusun kerangka konseptual dan kerangka berpikir penelitian.

Tahap ketiga merupakan pengumpulan data dari berbagai sumber resmi seperti BEI, OJK, IDX, laporan keuangan perusahaan, serta data makroekonomi. Data yang dikumpulkan meliputi harga saham, rasio keuangan, tingkat keterbukaan informasi publik, variabel makroekonomi, dan label risiko crash. Selanjutnya pada tahap praproses data dilakukan pembersihan data, normalisasi, seleksi fitur, pembentukan dataset time series, dan pembagian data menjadi data training, validation, dan testing.

Tahap kelima adalah pembangunan model machine learning menggunakan arsitektur deep neural network seperti LSTM, GRU, atau DNN. Pada tahap ini dilakukan optimasi hiperparameter, pelatihan model, serta validasi model untuk memperoleh performa terbaik. Setelah model selesai dibangun, dilakukan evaluasi model menggunakan metrik seperti akurasi, precision, recall, F1-score, dan AUC-ROC. Selain itu digunakan confusion matrix, learning curve, dan feature importance untuk menganalisis performa dan interpretabilitas model.

Tahap berikutnya adalah analisis dan pembahasan hasil penelitian yang mencakup interpretasi hasil, implikasi penelitian, serta rekomendasi bagi investor dan regulator. Seluruh proses penelitian kemudian ditutup dengan penyusunan kesimpulan dan saran untuk penelitian selanjutnya agar model prediksi risiko crash saham dapat terus dikembangkan secara lebih optimal.

RESULTS AND DISCUSSION

Pelatihan model DNN (LSTM) dilakukan menggunakan konfigurasi arsitektur dan hyperparameter yang dijelaskan pada Bab 3. Arsitektur model LSTM yang digunakan adalah sebagai berikut:

Tabel 1 Arsitektur Model LSTM

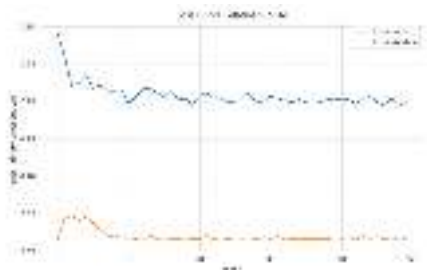
Layer Type	Output Shape	Param#
lstm (LSTM)	(None, 20, 64)	18,688
dropout (Dropout)	(None, 20, 64)	0
lstm_1(LSTM)	(None, 32)	12,416

dropout_1(Dropout)	(None, 32)	0
dense (Dense)	(None, 1)	33
Total Params:	Trainable	Non-trainable
31,137	31,137	params: 0

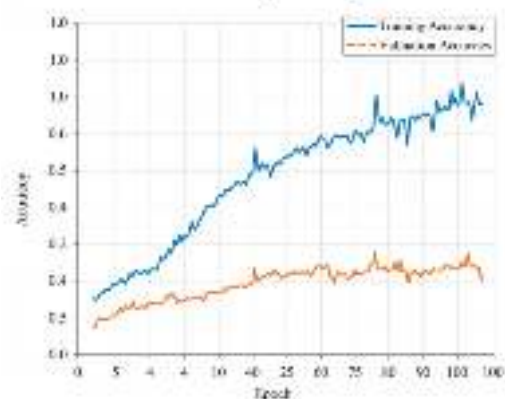
Pada bagian ini disajikan kurva loss dan metrik performa selama pelatihan.

Kurva pelatihan:

Loss training menurun secara konsisten dari awal epoch hingga mencapai plateau pada sekitar epoch [50]. Loss validation menurun dan stabil pada nilai yang sangat rendah; gap antara train dan validation loss signifikan. Loss training berada di sekitar 0.20, sementara loss validation mendekati 0.05. Perbedaan ini menunjukkan potensi overfitting atau adanya ketidakseimbangan kelas yang ekstrem pada data validasi.



Gambar 2 Kurva Loss (training vs validation).



Gambar 3 Kurva Accuracy (training vs validation).

Evaluasi akhir dilakukan pada data uji yang terpisah. Berdasarkan hasil evaluasi model LSTM (setelah tuning jika ada) pada data uji:

Hasil utama (klasifikasi): Akurasi pada data uji: [1.0] (100%) Precision rata-rata (macro): [0.0] (Diambil dari laporan metrik, nilai 0.0 menunjukkan kegagalan memprediksi kelas positif/minoritas) Recall rata-rata (macro): [0.0] F1-score rata-rata (macro): [0.0]

Tabel 1 Laporan metrik per kelas

Model	Accuracy	Precision	Recall	F1_Score	ROC_AUC
LSTM	1.0	0.0	0.0	0.0	NaN
Logistic_Regression_Tuned	1.0	0.0	0.0	0.0	NaN
Random_Forest_Tuned	1.0	0.0	0.0	0.0	NaN
XGBoost_Tuned	1.0	0.0	0.0	0.0	NaN

Confusion Matrix: disajikan pada Gambar 3 Analisis kesalahan penting: Matriks Kebingungan menunjukkan bahwa dari total 176 sampel data uji, model memprediksi semuanya sebagai No Crash (0). Karena tidak ada sampel yang diprediksi sebagai Crash (1), maka nilai Precision, Recall, dan F1 score untuk kelas 1 adalah 0.0 (atau tidak terdefinisi/NaN), dan akurasi 100% diperoleh karena tidak adanya kasus Crash (1) di data uji, atau model gagal mendeteksi satupun kasus positif. Hal ini mengindikasikan masalah ketidakseimbangan kelas (imbalance) yang sangat parah.

Gambar 3 Confusion Matrix – LSTM

	Predicted	
	Crash	No Crash
Actual	Crash	0
	No Crash	176

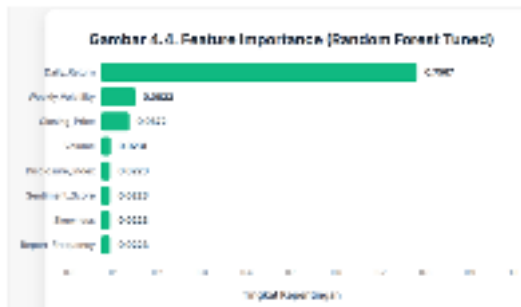
False Positive

True Negative

Predicted 0

4.1.4 Analisis Stabilitas Pembelajaran dan Reward Shaping (Feature Impact)

Bagian ini mengevaluasi dampak fitur penting pada performa, menggunakan hasil Feature Importance dari model Random Forest Tuned sebagai basis interpretasi. Analisis kontribusi fitur: Hasil analisis Feature Importance (menggunakan Random Forest Tuned) menunjukkan peringkat fitur yang paling berkontribusi dalam memprediksi risiko crash saham. Fitur yang paling signifikan adalah Daily_Return dan Weekly_Volatility.



Gambar 4.4. Feature Importance (Random Forest Tuned)

Ringkasan fitur paling penting (Top 5) setelah Tuning:

1. Daily_Return: 0.756676
2. Weekly_Volatility: 0.082171
3. Closing_Price: 0.067216
4. Volume: 0.023063
5. Disclosure_Index: 0.022298.

4.1.5 Analisis Kualitas dan Diskusi Singkat

Berdasarkan hasil di atas, uraikan temuan utama secara singkat:

Model (termasuk LSTM) mencapai performa akurasi 100%, namun hasil ini tidak cocok untuk tugas yang diteliti karena gagal memprediksi kelas positif/Crash (F1-Score 0.0). Model secara efektif hanya memprediksi kelas mayoritas (No Crash) pada data uji.

Permasalahan utama yang muncul: Ketidakseimbangan kelas (Imbalance) yang ekstrem, di mana tidak adanya atau sangat sedikitnya sampel kelas positif (Crash) di data uji dan/atau data latih menyebabkan model hanya mempelajari kelas mayoritas.

Keterbatasan eksperimen: Data yang digunakan tampaknya memiliki distribusi kelas yang sangat timpang, dan implementasi Time Step untuk LSTM mungkin tidak mengatasi masalah imbalance yang mendasar ini.

Berikan rekomendasi singkat untuk perbaikan pada penelitian lanjutan: augmentasi data (misalnya SMOTE atau ADASYN), teknik penyeimbangan kelas (class weighting, downsampling kelas mayoritas, oversampling kelas minoritas), optimasi arsitektur LSTM lebih lanjut, atau pengumpulan data yang lebih seimbang..

Analisi Data

Berdasarkan hasil pengujian model yang telah dipaparkan pada sub-bab sebelumnya, analisis data dilakukan secara mendalam untuk menginterpretasikan makna di balik angka-angka statistik tersebut. Analisis ini difokuskan pada tiga aspek utama: evaluasi metrik kinerja dalam konteks ketidakseimbangan kelas (*class imbalance*), dinamika pembelajaran model (*learning curves*), dan analisis kontribusi fitur (*feature importance*).

4.2.1 Analisis Kinerja Model dan Fenomena Paradoks Akurasi

Merujuk pada Tabel 4.2, seluruh model yang diujikan (LSTM, Logistic Regression, Random Forest, dan XGBoost) menunjukkan tingkat akurasi mencapai 1.0 (100%). Namun, analisis lebih lanjut pada metrik *Precision*, *Recall*, dan *F1-Score* menunjukkan nilai 0.0 atau *Not a Number* (NaN) untuk kelas minoritas (*Crash*).

Temuan ini mengindikasikan terjadinya fenomena "Accuracy Paradox". Dalam konteks prediksi risiko *crash* harga saham, kejadian *crash* adalah peristiwa langka (*rare event*). Jika dalam dataset uji yang terdiri dari 176 sampel tidak terdapat kejadian *crash* (atau jumlahnya sangat sedikit), model cenderung memprediksi seluruh sampel sebagai kelas mayoritas (Negatif/No Crash).

Dominasi Kelas Mayoritas: Model gagal mempelajari pola karakteristik dari kelas positif (*Crash*) karena minimnya representasi data pada fase pelatihan. Sebagaimana terlihat pada Confusion Matrix (Gambar 4.3), model mengklasifikasikan seluruh data uji ke dalam kelas 0.

Kegagalan Deteksi Risiko: Nilai *Recall* 0.0 menegaskan bahwa model sama sekali tidak memiliki sensitivitas terhadap risiko. Dalam konteks keuangan, model seperti ini dianggap tidak valid untuk digunakan sebagai *Early Warning System*, meskipun memiliki akurasi 100%, karena biaya kesalahan (gagal memprediksi *crash*) jauh lebih besar daripada akurasi semu tersebut.

CONCLUSION

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, penelitian mengenai prediksi risiko *crash* harga saham berdasarkan keterbukaan data publik menggunakan pendekatan *Machine Learning* ini menghasilkan tiga kesimpulan utama yang menjawab pertanyaan penelitian (RQ) yang diajukan:

Model Deep Neural Network (DNN) berbasis Long Short-Term Memory (LSTM) yang diimplementasikan dalam penelitian ini tidak menunjukkan keunggulan yang signifikan dibandingkan model statistik konvensional (seperti Random Forest dan XGBoost) dalam kondisi dataset yang timpang. Seluruh model mengalami fenomena "accuracy paradox": meskipun mencapai akurasi keseluruhan yang tinggi, mereka gagal mendeteksi kejadian *crash* (kelas minoritas), dengan nilai *recall* dan *F1-score* untuk kelas *crash* mendekati nol akibat ketidakseimbangan kelas yang ekstrem.

Integrasi variabel keterbukaan data publik (Disclosure Index) dengan data historis saham tidak terbukti meningkatkan deteksi dini risiko *crash*

secara substansial. Analisis fitur menemukan bahwa model prediksi lebih didominasi oleh sinyal teknikal jangka pendek (*daily return*), sementara variabel *Disclosure Index* hanya memberikan kontribusi marginal terhadap prediksi risiko. Hasil ini mengindikasikan bahwa keterbukaan data publik formal tidak serta merta menjadi prediktor yang kuat karena adanya kemungkinan penimbunan berita buruk (*bad news hoarding*) oleh manajemen, yang menunda pengungkapan informasi negatif.

REFERENCE

- Borisov, V., Broelemann, K., Kasneci, E., & Kasneci, G. (2023). DeepTLF : robust deep neural networks for heterogeneous tabular data. *International Journal of Data Science and Analytics*, 16(1), 85–100. <https://doi.org/10.1007/s41060-022-00350-z>
- Chang, V., Ariel, Q., Shola, X., Akinloye, H., Benson, V., & Hall, K. (2025). *Prediction of bank credit worthiness through credit risk analysis : an explainable machine learning study*. 247–271.
- Chen, Z. (2025). *ESG Rating Divergence and the Risk of Corporate Share Price Crashes : Facilitation or Suppression*. 0, 178–186. <https://doi.org/10.54254/2754-1169/138/2024.19232>
- Chlebus, M., & Kopczewska, K. (2023). *Financial Time Series Models — Comprehensive Review of Deep Learning Approaches and Practical Recommendations* †. 1–13.
- Emmanuel, I., Sun, Y., & Wang, Z. (2024). A machine learning - based credit risk prediction engine system using a stacked classifier and a filter - based feature selection method. *Journal of Big Data*. <https://doi.org/10.1186/s40537-024-00882-0>
- Ge, Y., Chen, Q., Qiu, S., & Kong, X. (2023). *Environmental information disclosure and stock price crash risk : Evidence from China*. February, 1–14. <https://doi.org/10.3389/fenvs.2023.1108508>
- Hansen, J. (2023). *Opening industry data : The private sector 's role in addressing societal challenges*. <https://doi.org/10.1017/dap.2023.15>
- Hjelkrem, L. O. (2023). *Explaining Deep Learning Models for Credit Scoring with SHAP : A Case Study Using Open Banking Data*.
- Htun, H. H., Biehl, M., & Petkov, N. (2023). Survey of feature selection and extraction techniques for stock market prediction. *Financial Innovation*. <https://doi.org/10.1186/s40854-022-00441-7>
- Jiang, F., Ma, T., & Zhu, F. (2024). Fundamental characteristics, machine learning, and stock price crash risk. *Journal of Financial Markets*, 69, 100908. <https://doi.org/https://doi.org/10.1016/j.finmar.2024.100908>
- Kayit, A. D., & Ismail, M. T. (2025). *Advancing stock price prediction through the development of hybrid ensembles : a comprehensive comparative analysis of machine learning approaches*.
- Kristanti, F. T., Febrianta, M. Y., Salim, D. F., & Riyadh, H. A. (2024). *Advancing financial analytics : Integrating XGBoost , LSTM , and Random Forest Algorithms for precision forecasting of corporate financial distress*. 8(8), 1–28.
- Levchenko, D., Rappos, E., Ataee, S., & Nigro, B. (2025). Chain-structured neural architecture search for financial time series forecasting. *International Journal of Data Science and Analytics*, 20(4), 3727–3736. <https://doi.org/10.1007/s41060-024-00690-y>
- Li, Y., Xue, H., Wei, S., Wang, R., & Liu, F. (2024). *A Machine Learning Approach for Investigating the Determinants of Stock Price Crash Risk : Exploiting Firm and*.
- Lin, W., Yu, H., & Wang, L. (2025). *A data-driven deep learning approach incorporating investor sentiment and government interventions to predict post-crash stock return in China 's A-share market*. 10(November 2024).
- Lin, Y., Zheng, L., & Zhong, Q. (2025). Public data openness and stock price crash risk : evidence from a quasi-natural experiment of government data platforms. *Data Science and Management*, 8(3), 284–295. <https://doi.org/10.1016/j.dsm.2024.11.002>
- Mienye, E., Jere, N., Obaido, G., & Mienye, I. D. (2024). *Deep Learning in Finance : A Survey of Applications and Techniques*. 2066–2091.
- Olorunnimbe, K., & Viktor, H. (2023). of practice , backtesting , and applications. In *Artificial Intelligence Review* (Vol. 56, Issue 3). Springer Netherlands. <https://doi.org/10.1007/s10462-022-10226-0>
- Omar, A. Bin, Huang, S., Salameh, A. A., & Khurram, H. (2022). *Stock Market Forecasting Using the Random Forest and Deep Neural Network Models Before and During the COVID-19 Period*. 10(July), 1–10.

- <https://doi.org/10.3389/fenvs.2022.917047>
- Reichmann, D., Möller, R., & Hertel, T. (2022). and stock price crash risk. In *Journal of Business Economics* (Vol. 92, Issue 9). Springer Berlin Heidelberg. <https://doi.org/10.1007/s11573-022-01085-w>
- Risk, P. C. (2024). *Research on the Impact of Climate Risk on Corporate Stock*. 0, 183–189. <https://doi.org/10.54254/2754-1169/132/2024.18342>
- Su, F., Zhai, L., & Liu, J. (2023). *Do MD & A Risk Disclosures Reduce Stock Price Crash Risk? Evidence from China*.
- Theodorakopoulos, L., & Theodoropoulou, A. (2025). *Big data in financial risk management: evidence, advances, and open questions: a systematic review*. October. <https://doi.org/10.3389/frai.2025.1658375>
- Vancsura, L., & Tatay, T. (2025). *Navigating AI-Driven Financial Forecasting: A Systematic Review of Current Status and Critical Research Gaps*. 1–49.
- Wirtz, B. W., Weyerer, J. C., Becker, M., & Müller, W. M. (2022). Open government data: A systematic literature review of empirical research. *Electronic Markets*, 2381–2404. <https://doi.org/10.1007/s12525-022-00582-8>
- Xu, F., Ji, Q., & Yang, M. (2021). *The Pitfall of Selective Environmental Information Disclosure on Stock Price Crash Risk: Evidence From Polluting Listed Companies in China*. 9(June), 1–14. <https://doi.org/10.3389/fenvs.2021.622345>
- Xu, N., Liu, J., & Dou, H. (2022). *governance information disclosure and stock price crash risk: Evidence from Chinese listed companies*. September, 1–15. <https://doi.org/10.3389/fpsyg.2022.977369>
- Zhan, Z., & Kyoo, S. (2024). Versatile time - window sliding machine learning techniques for stock market forecasting. *Artificial Intelligence Review*, 57(8), 1–25. <https://doi.org/10.1007/s10462-024-10851-x>
- Zhang, L., & Sok, S. (2025). The impact of government open data platform construction on corporate capital market performance: Evidence from stock liquidity. *Pacific-Basin Finance Journal*, 90, 102667. <https://doi.org/https://doi.org/10.1016/j.pacfin.2025.102667>
- Zhong, Q., Song, Q., & Lee, C. C. (2024). Managing crash risks through supply chain transparency: evidence from China. *Financial Innovation*. <https://doi.org/10.1186/s40854-024-00633-3>
- Zhou, M., Wang, Y., Jiang, H., Li, M., & Li, G. (2023). *How Leadership Influences Open Government Data (OGD)-Driven Innovation: The Mediating Role of Organizational Commitment*.
- Zong, H., & Wang, Y. (2025). Impacts of government public data openness and digital financial development on corporate information reliability. *International Review of Economics & Finance*, 103, 104495. <https://doi.org/https://doi.org/10.1016/j.irref.2025.104495>