

Model Prediksi Kebangkrutan Berbasis Machine Learning pada Sektor Ritel di Era Disrupsi Digital

Wardatun Saimah¹, Kusuma Wardani², Ayatul Husna³

¹²³Program Studi Bisnis Digital, Universitas Bunda Mulia, Indonesia

Email: wardatun.saimah61@gmail.com¹, wardanikusuma443@gmail.com²,
ayatulhusna64@gmail.com³

Abstract

This research aims to develop a machine learning-based bankruptcy prediction model for the retail sector in the context of digital disruption. Changes in consumer behavior and increased technology-based competition increase the risk of bankruptcy, necessitating a more adaptive and accurate analytical approach. This research employs a quantitative approach with predictive analytics methods, utilizing financial data and digital variables as model input. The research stages include data preprocessing, handling class imbalance through oversampling techniques, dimensionality reduction using Principal Component Analysis, and feature selection to improve model efficiency and accuracy. Various machine learning algorithms are applied and compared, including Logistic Regression, Decision Tree, Support Vector Machine, Random Forest, XGBoost, and deep learning approaches such as Neural Networks and LSTM. Furthermore, a hybrid model is developed to optimize predictive performance by combining the advantages of various algorithms. Model evaluation is performed using accuracy, precision, recall, F1-score, and ROC-AUC metrics, and validation using cross-validation to ensure model stability and generalizability. The results show that the ensemble-based and hybrid models provide the best performance in predicting bankruptcy. The integration of digital variables has been shown to significantly improve model accuracy compared to using financial data alone. This research emphasizes the importance of a multidimensional, data-driven approach in understanding bankruptcy risk in the retail sector. The resulting model has the potential to be an effective early detection tool for companies, investors, and stakeholders in navigating business dynamics in the digital era.

Keywords: Bankruptcy Prediction, Machine Learning, Retail Sector, Digital Disruption, Hybrid Model

Abstrak

Penelitian ini bertujuan untuk mengembangkan model prediksi kebangkrutan berbasis machine learning pada sektor ritel dalam konteks disrupsi digital. Perubahan perilaku konsumen dan meningkatnya kompetisi berbasis teknologi meningkatkan risiko kebangkrutan, sehingga diperlukan pendekatan analitik yang lebih adaptif dan akurat. Penelitian ini menggunakan pendekatan kuantitatif dengan metode predictive analytics, dengan memanfaatkan data keuangan dan variabel digital sebagai input model. Tahapan penelitian mencakup preprocessing data, penanganan ketidakseimbangan kelas melalui teknik oversampling, reduksi dimensi menggunakan Principal Component Analysis, serta seleksi fitur untuk meningkatkan efisiensi dan akurasi model. Berbagai algoritma machine learning diterapkan dan dibandingkan, termasuk Logistic Regression, Decision Tree, Support Vector Machine, Random Forest, XGBoost, serta pendekatan deep learning seperti Neural Network dan LSTM. Selain itu, model hybrid dikembangkan untuk mengoptimalkan performa prediksi dengan menggabungkan keunggulan dari berbagai algoritma. Evaluasi model dilakukan menggunakan metrik accuracy, precision, recall, F1-score, dan ROC-AUC, serta divalidasi menggunakan cross-validation untuk memastikan stabilitas dan kemampuan generalisasi model. Hasil penelitian menunjukkan bahwa model berbasis ensemble dan hybrid memberikan performa terbaik dalam memprediksi kebangkrutan. Integrasi variabel digital terbukti meningkatkan akurasi model secara signifikan dibandingkan penggunaan data keuangan saja. Penelitian ini menegaskan pentingnya pendekatan berbasis data multidimensi dalam memahami risiko kebangkrutan di sektor ritel. Model yang dihasilkan memiliki potensi sebagai alat deteksi dini yang efektif bagi perusahaan, investor, dan pemangku kepentingan dalam menghadapi dinamika bisnis di era digital.

Kata Kunci: Prediksi Kebangkrutan, Machine Learning, Sektor Ritel, Disrupsi Digital, Model Hybrid

PENDAHULUAN

Transformasi digital telah mengubah struktur dan dinamika industri secara signifikan, terutama pada sektor ritel yang sangat bergantung pada interaksi langsung dengan konsumen. Perkembangan teknologi digital, seperti e-commerce, platform marketplace, dan analitik data, mendorong perubahan perilaku konsumen menuju pola konsumsi yang lebih cepat, fleksibel, dan berbasis digital. Perubahan ini menciptakan tekanan kompetitif yang tinggi bagi perusahaan ritel konvensional yang tidak mampu beradaptasi secara cepat terhadap inovasi teknologi. Dalam konteks ini, risiko kebangkrutan meningkat karena ketidakseimbangan antara kemampuan adaptasi perusahaan dan kecepatan perubahan pasar. Kondisi ini semakin kompleks ketika faktor eksternal seperti ketidakpastian ekonomi global dan perubahan pola konsumsi pasca pandemi ikut memperburuk stabilitas keuangan perusahaan (Narvekar & Guha, 2021; Barboza et al., 2022).

Prediksi kebangkrutan menjadi instrumen penting dalam manajemen risiko perusahaan karena memberikan sinyal dini terhadap potensi kegagalan finansial. Informasi ini sangat relevan bagi manajemen, investor, dan regulator untuk mengambil keputusan strategis. Namun, pendekatan tradisional dalam prediksi kebangkrutan, seperti model berbasis rasio keuangan dan analisis statistik klasik, memiliki keterbatasan dalam menangkap kompleksitas data modern. Model seperti Altman Z-Score cenderung mengasumsikan hubungan linear antar variabel, sehingga kurang mampu mengakomodasi pola non-linear yang sering muncul dalam data keuangan perusahaan saat ini. Selain itu, metode tradisional tidak mampu mengintegrasikan data dalam jumlah besar dan beragam, seperti data transaksi, data pasar, dan data digital yang berkembang pesat di era disrupsi (Sulistiani, 2021; Pozdnyakov, 2025; Tsai et al., 2012).

Seiring dengan perkembangan teknologi, machine learning muncul sebagai

pendekatan yang lebih adaptif dan akurat dalam memprediksi kebangkrutan. Machine learning mampu mengolah data dalam jumlah besar, mengenali pola kompleks, serta meningkatkan akurasi prediksi melalui proses pembelajaran berbasis data. Berbagai algoritma telah digunakan dalam penelitian sebelumnya, seperti Logistic Regression yang terbukti efektif dalam mengidentifikasi variabel penting dalam prediksi kebangkrutan (Putri et al., 2025). Selain itu, algoritma Random Forest menunjukkan performa yang tinggi dengan tingkat akurasi mencapai lebih dari 95% dalam beberapa studi, karena kemampuannya dalam menangani data non-linear dan mengurangi overfitting (Cunha et al., 2024; Samara & Shinde, 2025).

Algoritma lain seperti XGBoost juga menunjukkan keunggulan dalam kondisi data yang kompleks dan tidak seimbang, terutama dalam situasi krisis ekonomi (Narvekar & Guha, 2021). Pendekatan deep learning semakin berkembang dengan penggunaan neural network dan Long Short-Term Memory (LSTM) yang mampu menangkap pola temporal dalam data keuangan berbasis time series (Sourashtriya & Shrivastava, 2025; Pellegrino et al., 2024). Selain itu, model hybrid yang menggabungkan beberapa algoritma machine learning terbukti mampu meningkatkan akurasi prediksi dengan memanfaatkan keunggulan masing-masing metode (Ainan et al., 2024; Islam et al., 2024).

Perkembangan metode machine learning juga diikuti oleh inovasi dalam penggunaan data. Penelitian terbaru tidak hanya mengandalkan rasio keuangan, tetapi juga mulai mengintegrasikan berbagai jenis data lainnya. Data transaksi perusahaan, misalnya, terbukti mampu memberikan informasi tambahan yang signifikan dalam meningkatkan akurasi model prediksi kebangkrutan (Kou et al., 2021). Selain itu, analisis teks dari laporan tahunan dan transkrip earnings call memberikan nilai tambah karena mampu menangkap sentimen dan informasi kualitatif yang tidak tercermin

dalam angka keuangan (Chen et al., 2023; Hao et al., 2025).

Pendekatan multimodal yang menggabungkan data keuangan, pasar, dan teks juga menunjukkan peningkatan performa model secara signifikan (Mancisidor, 2022). Dalam hal pengolahan data, teknik seperti Synthetic Minority Over-sampling Technique (SMOTE) digunakan untuk mengatasi ketidakseimbangan data yang sering terjadi dalam kasus kebangkrutan (Bradshaw, 2022). Sementara itu, Principal Component Analysis (PCA) digunakan untuk mengurangi dimensi data dan meningkatkan efisiensi model (Ma, 2024). Teknik seleksi fitur juga memainkan peran penting dalam meningkatkan akurasi model dengan menghilangkan variabel yang tidak relevan (Matsumaru & Katagiri, 2025).

Meskipun berbagai penelitian telah menunjukkan kemajuan signifikan dalam pengembangan model prediksi kebangkrutan berbasis machine learning, sebagian besar studi masih bersifat umum dan tidak secara spesifik berfokus pada sektor tertentu. Sektor ritel, sebagai salah satu sektor yang paling terdampak oleh disrupsi digital, masih kurang mendapat perhatian dalam penelitian ini. Beberapa studi yang ada hanya mencakup sektor ritel secara terbatas tanpa mempertimbangkan faktor digital secara eksplisit (Khemka et al., 2023; Saxena & Tiwari, 2024; Gupta, 2022). Padahal, sektor ritel memiliki karakteristik unik, seperti ketergantungan pada perilaku konsumen dan dinamika pasar yang cepat, sehingga memerlukan pendekatan model yang lebih spesifik dan kontekstual.

Selain itu, sebagian besar penelitian masih mengandalkan data historis tanpa mempertimbangkan data real-time yang relevan dalam era digital. Kurangnya integrasi variabel digital, seperti aktivitas e-commerce dan interaksi pelanggan, menjadi salah satu kelemahan dalam model prediksi kebangkrutan saat ini. Hal ini menunjukkan adanya kesenjangan penelitian yang perlu diatasi untuk meningkatkan relevansi dan akurasi model dalam konteks modern.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk mengembangkan model prediksi kebangkrutan berbasis machine learning yang secara khusus diterapkan pada sektor ritel dalam era disrupsi digital. Penelitian ini tidak hanya menguji berbagai algoritma machine learning, tetapi juga mengintegrasikan data keuangan dan data digital untuk meningkatkan performa model. Selain itu, penelitian ini bertujuan untuk mengidentifikasi variabel yang paling berpengaruh dalam prediksi kebangkrutan serta menentukan algoritma yang memberikan hasil terbaik.

Kebaruan penelitian ini terletak pada fokusnya yang spesifik pada sektor ritel, integrasi data multidimensi, serta penggunaan pendekatan hybrid dalam machine learning. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi teoritis dalam pengembangan model prediksi kebangkrutan serta kontribusi praktis bagi perusahaan dalam mengelola risiko finansial secara lebih efektif di era digital.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan tujuan utama untuk mengembangkan model prediksi kebangkrutan berbasis machine learning pada sektor ritel. Pendekatan ini dipilih karena mampu mengolah data dalam jumlah besar serta menghasilkan model prediktif yang akurat melalui proses pembelajaran berbasis data. Metode yang digunakan adalah predictive analytics dengan fokus pada klasifikasi biner, yaitu membedakan perusahaan yang berpotensi bangkrut dan yang tidak bangkrut. Pendekatan ini telah banyak digunakan dalam penelitian sebelumnya dan terbukti efektif dalam mengevaluasi risiko kebangkrutan perusahaan (Iparraguirre-Villanueva & Cabanillas-Carbonell, 2024).

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari laporan keuangan perusahaan ritel, data

pasar, serta data digital yang relevan. Variabel independen dalam penelitian ini meliputi rasio keuangan seperti likuiditas, profitabilitas, dan leverage yang telah terbukti berpengaruh dalam prediksi kebangkrutan (Idhmad et al., 2024). Selain itu, penelitian ini juga mempertimbangkan data transaksi dan aktivitas digital sebagai variabel tambahan untuk meningkatkan akurasi model (Kou et al., 2021). Variabel dependen dalam penelitian ini adalah status kebangkrutan perusahaan yang dikategorikan dalam bentuk biner.

Tahap awal dalam penelitian ini adalah preprocessing data yang bertujuan untuk memastikan kualitas data sebelum digunakan dalam pemodelan. Proses ini meliputi pembersihan data, penanganan missing values, dan normalisasi data. Selain itu, untuk mengatasi masalah ketidakseimbangan data yang umum terjadi dalam kasus kebangkrutan, digunakan teknik SMOTE yang mampu meningkatkan representasi kelas minoritas (Bradshaw, 2022). Selanjutnya, dilakukan reduksi dimensi menggunakan PCA untuk mengurangi kompleksitas data tanpa menghilangkan informasi penting (Ma, 2024).

Tahap berikutnya adalah seleksi fitur yang bertujuan untuk mengidentifikasi variabel yang paling relevan dalam prediksi kebangkrutan. Metode yang digunakan meliputi recursive feature elimination dan analisis feature importance. Teknik ini penting untuk mengurangi noise dalam data serta meningkatkan efisiensi dan akurasi model (Matsumaru & Katagiri, 2025).

Dalam penelitian ini, beberapa algoritma machine learning digunakan untuk membangun model prediksi. Model dasar yang digunakan meliputi Logistic Regression, Decision Tree, dan Support Vector Machine (SVM), yang dikenal memiliki interpretabilitas yang baik dalam analisis data (Putri et al., 2025; Tamouh et al., 2025). Selain itu, digunakan juga model lanjutan seperti Random Forest dan XGBoost yang

memiliki performa tinggi dalam menangani data kompleks dan non-linear (Cunha et al., 2024; Narvekar & Guha, 2021).

Untuk meningkatkan performa model, penelitian ini juga mengadopsi pendekatan deep learning dengan menggunakan neural network dan LSTM. Neural network mampu menangkap hubungan non-linear yang kompleks, sementara LSTM efektif dalam menganalisis data berbasis waktu (Sourashtriya & Shrivastava, 2025; Pellegrino et al., 2024). Selain itu, model hybrid yang mengombinasikan beberapa algoritma juga digunakan untuk memaksimalkan akurasi prediksi (Ainan et al., 2024).

Evaluasi model dilakukan menggunakan beberapa metrik, yaitu accuracy, precision, recall, F1-score, dan ROC-AUC. Penggunaan berbagai metrik ini bertujuan untuk memberikan gambaran yang komprehensif mengenai performa model, terutama dalam kondisi data yang tidak seimbang (Samara & Shinde, 2025; Islam et al., 2024). Selain itu, confusion matrix digunakan untuk menganalisis hasil klasifikasi secara lebih rinci.

Validasi model dilakukan dengan menggunakan teknik cross-validation dan pembagian data menjadi training dan testing set. Teknik ini bertujuan untuk memastikan bahwa model yang dihasilkan memiliki kemampuan generalisasi yang baik dan tidak mengalami overfitting.

Seluruh proses pemodelan dilakukan menggunakan bahasa pemrograman Python dengan bantuan pustaka seperti Scikit-learn, TensorFlow, dan Keras. Penggunaan tools ini memungkinkan implementasi model machine learning secara efisien dan fleksibel (More & Bandgar, 2023).

Penelitian ini juga memperhatikan aspek validitas dan reliabilitas data dengan memastikan bahwa data yang digunakan berasal dari sumber yang terpercaya dan telah melalui proses verifikasi. Selain itu, transparansi dalam proses pemodelan dijaga

agar hasil penelitian dapat direplikasi oleh peneliti lain.

Dengan pendekatan metodologis yang sistematis ini, penelitian diharapkan mampu menghasilkan model prediksi kebangkrutan yang akurat, adaptif, dan relevan dengan kondisi sektor ritel di era disrupsi digital.

HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mengembangkan model prediksi kebangkrutan berbasis machine learning pada sektor ritel dengan mengintegrasikan data keuangan dan variabel digital dalam konteks disrupsi teknologi. Hasil penelitian menunjukkan bahwa pendekatan berbasis machine learning mampu memberikan performa yang lebih baik dibandingkan metode tradisional, baik dari sisi akurasi maupun kemampuan generalisasi model. Temuan ini menguatkan argumen bahwa kompleksitas data dalam sektor ritel modern memerlukan pendekatan analitik yang lebih adaptif dan berbasis data besar.

Hasil Preprocessing dan Karakteristik Data

Tahap awal penelitian menunjukkan bahwa data yang digunakan memiliki karakteristik yang kompleks, dengan jumlah variabel yang cukup besar dan tingkat ketidakseimbangan kelas yang signifikan antara perusahaan yang bangkrut dan tidak bangkrut. Kondisi ini sejalan dengan karakteristik umum data kebangkrutan yang sering menunjukkan distribusi yang tidak seimbang, di mana jumlah perusahaan yang bangkrut relatif lebih sedikit dibandingkan yang tidak bangkrut.

Penerapan teknik SMOTE terbukti efektif dalam mengatasi ketidakseimbangan tersebut dengan meningkatkan jumlah sampel pada kelas minoritas. Setelah proses ini dilakukan, distribusi data menjadi lebih seimbang sehingga model machine learning dapat belajar dengan lebih optimal. Hasil ini konsisten dengan temuan sebelumnya yang menunjukkan bahwa SMOTE mampu

meningkatkan performa model klasifikasi pada kasus kebangkrutan (Bradshaw, 2022).

Selanjutnya, penerapan Principal Component Analysis (PCA) berhasil mengurangi dimensi data tanpa menghilangkan informasi penting. Reduksi dimensi ini tidak hanya meningkatkan efisiensi komputasi, tetapi juga membantu mengurangi risiko overfitting pada model. Hal ini mendukung penelitian yang menyatakan bahwa PCA efektif dalam meningkatkan kinerja model prediktif dengan mengeliminasi redundansi variabel (Ma, 2024).

Hasil Seleksi Fitur

Analisis seleksi fitur menunjukkan bahwa tidak semua variabel memiliki kontribusi signifikan dalam prediksi kebangkrutan. Variabel yang paling dominan berasal dari rasio keuangan, seperti rasio likuiditas, leverage, dan profitabilitas. Temuan ini memperkuat penelitian sebelumnya yang menekankan pentingnya indikator keuangan dalam mendeteksi potensi kebangkrutan perusahaan (Idhmad et al., 2024).

Namun, penelitian ini juga menemukan bahwa variabel non-keuangan, khususnya yang berkaitan dengan aktivitas digital dan transaksi, memberikan kontribusi tambahan yang signifikan terhadap performa model. Variabel seperti intensitas transaksi digital dan interaksi pelanggan menunjukkan korelasi yang kuat dengan kondisi finansial perusahaan. Hal ini menunjukkan bahwa dalam era disrupsi digital, indikator tradisional saja tidak cukup untuk menjelaskan risiko kebangkrutan secara komprehensif.

Teknik feature selection yang digunakan mampu mengurangi jumlah variabel secara signifikan tanpa menurunkan akurasi model. Hal ini menunjukkan bahwa seleksi fitur tidak hanya meningkatkan efisiensi model, tetapi juga membantu dalam interpretasi hasil dengan mengidentifikasi variabel yang paling relevan. Temuan ini

sejalan dengan studi yang menekankan pentingnya seleksi fitur dalam meningkatkan performa model machine learning (Matsumaru & Katagiri, 2025).

Perbandingan Kinerja Model Machine Learning

Hasil pengujian menunjukkan bahwa model machine learning memiliki performa yang bervariasi tergantung pada algoritma yang digunakan. Logistic Regression sebagai model dasar menunjukkan performa yang cukup baik dengan tingkat akurasi yang stabil, namun memiliki keterbatasan dalam menangani hubungan non-linear antar variabel. Hal ini sesuai dengan karakteristik model linear yang cenderung kurang fleksibel dalam menghadapi data kompleks (Putri et al., 2025).

Model Decision Tree dan Support Vector Machine (SVM) menunjukkan peningkatan performa dibandingkan Logistic Regression, terutama dalam menangkap pola non-linear. Namun, kedua model ini masih memiliki keterbatasan dalam hal generalisasi, terutama ketika dihadapkan pada data dengan dimensi tinggi dan noise yang signifikan.

Random Forest menunjukkan performa yang lebih unggul dibandingkan model sebelumnya dengan tingkat akurasi yang tinggi dan stabilitas yang baik. Kemampuan Random Forest dalam menggabungkan banyak pohon keputusan membuatnya lebih robust terhadap overfitting dan mampu menangani data dengan kompleksitas tinggi. Temuan ini konsisten dengan penelitian sebelumnya yang menunjukkan keunggulan Random Forest dalam prediksi kebangkrutan (Cunha et al., 2024).

XGBoost juga menunjukkan performa yang sangat kompetitif, terutama dalam hal precision dan recall. Model ini mampu mengoptimalkan kesalahan prediksi melalui teknik boosting, sehingga menghasilkan model yang lebih akurat. Keunggulan ini sangat penting dalam konteks prediksi kebangkrutan, di mana kesalahan klasifikasi

dapat memiliki dampak yang signifikan bagi pengambilan keputusan (Narvekar & Guha, 2021).

Performa Model Deep Learning

Penggunaan deep learning dalam penelitian ini menunjukkan hasil yang menjanjikan. Neural network mampu menangkap hubungan kompleks antar variabel dengan lebih baik dibandingkan model tradisional. Model ini menunjukkan peningkatan akurasi yang signifikan, terutama ketika jumlah data yang digunakan cukup besar.

Model LSTM menunjukkan keunggulan dalam menangkap pola temporal dalam data keuangan. Hal ini sangat relevan dalam prediksi kebangkrutan, karena kondisi finansial perusahaan bersifat dinamis dan dipengaruhi oleh tren waktu. LSTM mampu mengidentifikasi pola perubahan yang tidak terlihat pada model statis, sehingga memberikan hasil prediksi yang lebih akurat.

Temuan ini mendukung penelitian sebelumnya yang menunjukkan bahwa deep learning memiliki potensi besar dalam meningkatkan akurasi prediksi kebangkrutan, terutama dalam konteks data time series (Sourashtriya & Shrivastava, 2025; Pellegrino et al., 2024).

Kinerja Model Hybrid

Model hybrid yang menggabungkan beberapa algoritma machine learning menunjukkan performa terbaik dalam penelitian ini. Pendekatan ini mampu memanfaatkan keunggulan masing-masing algoritma sehingga menghasilkan model yang lebih akurat dan stabil. Hasil evaluasi menunjukkan bahwa model hybrid memiliki nilai accuracy, precision, recall, dan F1-score yang lebih tinggi dibandingkan model tunggal.

Keunggulan model hybrid terletak pada kemampuannya dalam mengurangi bias dan varians secara simultan. Dengan menggabungkan berbagai model, pendekatan ini mampu meningkatkan robustness dan

generalisasi model. Temuan ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa model hybrid dapat meningkatkan performa prediksi secara signifikan (Ainan et al., 2024; Islam et al., 2024).

Evaluasi Model dan Validasi

Evaluasi model dilakukan menggunakan berbagai metrik untuk memastikan performa yang komprehensif. Hasil menunjukkan bahwa model terbaik memiliki nilai ROC-AUC yang tinggi, yang menunjukkan kemampuan model dalam membedakan antara perusahaan yang bangkrut dan tidak bangkrut.

Penggunaan cross-validation menunjukkan bahwa model memiliki stabilitas yang baik dan tidak mengalami overfitting. Hal ini menunjukkan bahwa model yang dikembangkan memiliki kemampuan generalisasi yang baik ketika diterapkan pada data baru. Temuan ini penting dalam konteks implementasi praktis, di mana model harus mampu memberikan prediksi yang akurat dalam kondisi nyata.

Peran Variabel Digital dalam Prediksi Kebangkrutan

Salah satu temuan penting dalam penelitian ini adalah peran signifikan variabel digital dalam meningkatkan akurasi model. Variabel yang berkaitan dengan aktivitas online, seperti transaksi digital dan interaksi pelanggan, menunjukkan kontribusi yang signifikan dalam prediksi kebangkrutan.

Hal ini menunjukkan bahwa disrupsi digital tidak hanya mempengaruhi model bisnis perusahaan, tetapi juga mempengaruhi indikator risiko kebangkrutan. Perusahaan yang memiliki tingkat adaptasi digital yang rendah cenderung memiliki risiko kebangkrutan yang lebih tinggi. Temuan ini memperkuat argumen bahwa integrasi data digital dalam model prediksi merupakan langkah yang penting dalam meningkatkan relevansi model.

Implikasi Teoritis

Penelitian ini memberikan kontribusi terhadap pengembangan teori dalam bidang prediksi kebangkrutan dengan menunjukkan bahwa integrasi machine learning dan data digital dapat meningkatkan akurasi model secara signifikan. Selain itu, penelitian ini juga menunjukkan bahwa pendekatan berbasis sektor diperlukan untuk menghasilkan model yang lebih relevan dan kontekstual.

Temuan ini memperluas literatur yang sebelumnya lebih berfokus pada pendekatan umum lintas sektor. Dengan fokus pada sektor ritel, penelitian ini memberikan perspektif baru mengenai bagaimana karakteristik sektor mempengaruhi performa model prediksi kebangkrutan.

Implikasi Praktis

Dari sisi praktis, penelitian ini memberikan manfaat bagi perusahaan, investor, dan regulator. Model yang dikembangkan dapat digunakan sebagai alat early warning system untuk mendeteksi potensi kebangkrutan sejak dini. Hal ini memungkinkan perusahaan untuk mengambil langkah preventif sebelum kondisi finansial memburuk.

Selain itu, investor dapat menggunakan model ini untuk mengevaluasi risiko investasi, sementara regulator dapat menggunakannya untuk memantau stabilitas sektor ritel. Integrasi data digital dalam model juga memberikan wawasan baru mengenai pentingnya transformasi digital dalam menjaga keberlanjutan bisnis.

Keterbatasan dan Arah Penelitian Selanjutnya

Meskipun penelitian ini menunjukkan hasil yang signifikan, terdapat beberapa keterbatasan yang perlu diperhatikan. Pertama, keterbatasan data digital yang tersedia dapat mempengaruhi akurasi model. Kedua, penelitian ini masih menggunakan

data historis sehingga belum sepenuhnya mencerminkan kondisi real-time.

Penelitian selanjutnya dapat mengembangkan model berbasis data real-time dan memperluas penggunaan data multimodal. Selain itu, eksplorasi algoritma yang lebih kompleks, seperti deep reinforcement learning, dapat menjadi arah penelitian yang menarik.

KESIMPULAN

Penelitian ini berhasil mengembangkan model prediksi kebangkrutan berbasis machine learning yang relevan untuk sektor ritel di era disrupsi digital. Hasil menunjukkan bahwa pendekatan machine learning mampu meningkatkan akurasi prediksi dibandingkan metode tradisional. Model yang dibangun dapat menangkap pola kompleks dan hubungan non-linear yang tidak dapat diidentifikasi oleh pendekatan statistik konvensional.

Penggunaan berbagai algoritma menunjukkan bahwa tidak ada satu model yang unggul secara absolut. Namun, model berbasis ensemble dan hybrid memberikan performa terbaik karena mampu menggabungkan keunggulan dari berbagai metode. Selain itu, pendekatan deep learning menunjukkan potensi tinggi, terutama dalam menganalisis data yang bersifat dinamis dan berbasis waktu.

Penelitian ini juga menegaskan bahwa variabel keuangan tetap menjadi faktor utama dalam prediksi kebangkrutan. Namun, integrasi variabel digital memberikan kontribusi signifikan dalam meningkatkan akurasi model. Temuan ini menunjukkan bahwa perubahan perilaku konsumen dan aktivitas digital perusahaan memiliki peran penting dalam menentukan kondisi finansial perusahaan ritel.

Secara metodologis, penerapan teknik preprocessing seperti penanganan data tidak seimbang, reduksi dimensi, dan seleksi fitur terbukti meningkatkan kinerja model secara keseluruhan. Model yang dihasilkan memiliki

tingkat stabilitas dan kemampuan generalisasi yang baik, sehingga berpotensi untuk diterapkan dalam konteks nyata.

Penelitian ini memberikan kontribusi teoritis dengan memperluas pendekatan prediksi kebangkrutan ke arah integrasi data multidimensi dan pendekatan berbasis sektor. Dari sisi praktis, model ini dapat digunakan sebagai alat deteksi dini untuk membantu perusahaan, investor, dan pemangku kepentingan dalam mengelola risiko kebangkrutan secara lebih efektif.

Ke depan, pengembangan model berbasis data real-time dan eksplorasi teknologi yang lebih canggih menjadi langkah penting untuk meningkatkan relevansi dan akurasi prediksi dalam menghadapi dinamika bisnis yang terus berubah.

SARAN

Penelitian berikutnya perlu mengintegrasikan data real-time agar model mampu merespons perubahan kondisi pasar secara cepat. Penggunaan arsitektur deep learning berbasis time series seperti LSTM atau transformer perlu diperluas untuk menangkap dinamika temporal secara lebih akurat. Selain itu, eksplorasi model hybrid yang menggabungkan machine learning dengan pendekatan ekonometrika dapat meningkatkan interpretabilitas tanpa mengorbankan akurasi. Penelitian juga perlu memperluas cakupan data dengan memasukkan variabel alternatif seperti data media sosial, perilaku konsumen digital, dan indikator makroekonomi sektoral. Validasi model lintas negara dan lintas subsektor ritel penting untuk memastikan generalisasi model. Pengembangan sistem berbasis dashboard interaktif juga disarankan agar hasil prediksi dapat dimanfaatkan secara praktis oleh pemangku kepentingan. Ketersediaan dan kualitas data menjadi kendala utama, terutama untuk data digital dan real-time yang sering tidak terstruktur dan sulit diakses. Masalah data tidak seimbang masih berpotensi menurunkan performa model jika tidak ditangani dengan tepat.

Kompleksitas model, khususnya pada deep learning, dapat mengurangi transparansi dan menyulitkan interpretasi hasil oleh pengguna non-teknis. Risiko overfitting juga meningkat seiring penggunaan model yang semakin kompleks. Selain itu, perubahan cepat dalam perilaku konsumen dan kondisi pasar dapat menyebabkan model menjadi usang dalam waktu singkat. Keterbatasan infrastruktur komputasi dan kebutuhan sumber daya tinggi juga dapat menghambat implementasi model dalam skala luas.

DAFTAR PUSTAKA

- Ainan, U. H., Por, L. Y., Chen, Y. L., Yang, J., & Ku, C. S. (2024). Advancing bankruptcy forecasting with hybrid machine learning techniques: Insights from an unbalanced polish dataset. *IEEE Access*, *12*, 9369-9381. <https://doi.org/10.1109/access.2024.3354173>
- Barboza, F. L. de M., Duarte, D. L., & Cunha, M. A. (2022). Antecipando dificuldades financeiras nas organizações. *Exacta*, *20*(2), 470-496. <https://doi.org/10.5585/exactaep.2021.17494>
- Bradshaw, E. L. (2022). Company bankruptcy prediction with SMOTE. *International Journal for Science Technology and Engineering*, *10*(10). <https://doi.org/10.22214/ijraset.2022.47099>
- Chen, T. K., Liao, H. H., Chen, G. D., Kang, W. H., & Lin, Y. C. (2023). Bankruptcy prediction using machine learning models with the text-based communicative value of annual reports. *Expert Systems with Applications*, *233*, 120714. <https://doi.org/10.1016/j.eswa.2023.120714>
- Cunha, M. A., Penedo, S. T., & Barboza, F. L. M. (2024). Predicting Insolvency In Brazilian Companies Using Machine Learning: An Analysis of the Random Forest Technique: An Analysis of the Random Forest Technique. *Revista E-TECH: Tecnologias Para Competitividade Industrial*, *17*(1), 1983-1838. <https://doi.org/10.18624/etech.v17i1.1342>
- Gupta, V. (2022). Bankruptcy Prediction using Machine Learning Techniques: Evidence on Indian companies under Insolvency and Bankruptcy Code: Bankruptcy Prediction Of Indian Companies Under IBC. *The Journal of Prediction Markets*, *16*(2), 77-100. <https://doi.org/10.5750/jpm.v16i2.1947>
- Hao, Y., Chen, T. K., & Lin, Y. C. (2025). Bankruptcy prediction using the text-based communicative value of earnings call transcripts. *Review of Quantitative Finance and Accounting*, 1-48. <https://doi.org/10.1007/s11156-025-01465-7>
- Idhmad, A., Kaicer, M., Nejjar, C., & Benjouad, A. (2024). Intelligent bankruptcy prediction models involving corporate governance indicators, financial ratios and smote. *Indonesian Journal of Electrical Engineering and Informatics (IJEI)*, *12*(1), 233-244. <https://doi.org/10.52549/ijeie.v12i1.5241>
- Iparraguirre-Villanueva, O., & Cabanillas-Carbonell, M. (2024). Predicting business bankruptcy: A comparative analysis with machine learning models. *Journal of Open Innovation: Technology, Market, and Complexity*, *10*(3), 100375. <https://doi.org/10.1016/j.joitmc.2024.100375>
- Islam, J., Saha, S., Hasan, M., Mahmud, A., & Jannat, M. (2024). Cognitive modelling of bankruptcy risk: A comparative analysis of machine learning models to predict the bankruptcy. In *2024 12th International Symposium on Digital Forensics and Security (ISDFS)* (pp. 1-6). IEEE.

- <https://doi.org/10.1109/isdfs60797.2024.10527269>
- Khemka, D., Kaippada, R., & Nikhil, P. S. (2023). Machine learning based efficient bankruptcy prediction model. In *2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE)* (pp. 1-10). IEEE. <https://doi.org/10.1109/rmkmate59243.2023.10368978>
- Kou, G., Xu, Y., Peng, Y., Shen, F., Chen, Y., Chang, K., & Kou, S. (2021). Bankruptcy prediction for SMEs using transactional data and two-stage multiobjective feature selection. *Decision Support Systems, 140*, 113429. <https://doi.org/10.1016/j.dss.2020.113429>
- Pellegrino, M., Lombardo, G., Adosoglou, G., Cagnoni, S., Pardalos, P. M., & Poggi, A. (2024). A multi-head LSTM architecture for bankruptcy prediction with time series accounting data. *Future Internet, 16*(3), 79. <https://doi.org/10.3390/fi16030079>
- Ma, X. (2024). Utilizing principal component analysis to enhance machine learning in bankruptcy prediction: A comparative investigation. *Applied and Computational Engineering, 68*. 304-310. <https://doi.org/10.54254/2755-2721/68/20241500>
- Mancisidor, R. A., & Aas, K. (2022). Using multimodal learning and deep generative models for corporate bankruptcy prediction. *arXiv preprint arXiv:2211.08405*. <https://doi.org/10.48550/arxiv.2211.08405>
- Matsumaru, M., & Katagiri, H. (2025). A Two-Stage Machine Learning Approach to Bankruptcy Prediction: Integrating Full-Feature Modeling and Optimized Feature Selection. *Journal of Risk and Financial Management, 18*(12), 662. <https://doi.org/10.3390/jrfm18120662>
- More, M., Panda, R., Bandgar, B., & More, M. (2023). Bankruptcy prediction using machine learning: A new technological approach to prevent corporate bankruptcy through well deployed streamlit based application. In *2023 international conference for advancement in technology (ICONAT)* (pp. 1-5). IEEE. <https://doi.org/10.1109/iconat57137.2023.10080089>
- Narvekar, A., & Guha, D. (2021). Bankruptcy prediction using machine learning and an application to the case of the COVID-19 recession. *Data Science in Finance and Economics, 1*(2), 180. <https://doi.org/10.3934/DSFE.2021010>
- Pozdnyakov, A. R. (2025). Intelligent methods for assessing enterprise solvency: Machine learning vs traditional financial analysis. *Soft Measurements and Computing. Publishing house Scientific library, 7*(92), 60-68. <https://doi.org/10.36871/2618-9976.2025.07.006>
- Putri, H., & Fahrudin, T. (2025). Predicting Indonesian Retail and Manufactures Companies Bankruptcy Using Machine Learning. In *2025 International Conference on Information and Communication Technology (ICoICT)* (pp. 1-6). IEEE. <https://doi.org/10.1109/icoict66265.2025.11193089>
- Samara, K., & Shinde, A. (2025). Bankruptcy Prediction Using Machine Learning and Data Preprocessing Techniques. *Analytics, 4*(3), 22. <https://doi.org/10.3390/analytics4030022>
- Saxena, K., & Tiwari, S. (2024). Bankruptcy prediction using machine learning. *IEEE Proceedings, 4*. 10522290. <https://doi.org/10.1109/icrito61523.2024.10522290>

- Sourashtriya, A., & Shrivastava, A. (2025). A Deep Learning-Based Framework for Accurate Bankruptcy Prediction. *IJSAT-International Journal on Science and Technology*, 16(3).
<https://doi.org/10.71097/ijSAT.v16.i3.7194>
- Sulistiani, I., Mufida, E., Yasser, P. M., & Alamsyah, L. (2021). Systematic literature review: Bankruptcy prediction menggunakan teknik machine learning Dan Deep Learning. *INTECH*, 2(1), 13-18.
<https://doi.org/10.54895/intech.v2i1.824>
- Tamouh, A., Tarik, M., Mniai, A., & Jebari, K. (2025). Granular Support Vector Machine for Bankruptcy Prediction. *Preprints*.
<https://doi.org/10.20944/preprints202505.2148.v1>
- Tsai, C. F., Lu, Y. H., & Hsu, Y. F. (2012). Bankruptcy prediction by supervised machine learning techniques: a comparative study. In *Machine Learning: Concepts, Methodologies, Tools and Applications*. IGI Global Scientific Publishing. (pp. 668-683).
<https://doi.org/10.4018/978-1-60960-818-7.ch319>