

OPTIMALISASI AKURASI PREDIKSI PERSETUJUAN PINJAMAN UANG MENGUNAKAN ALGORITMA RANDOM FOREST

Ade Irwan¹, Rudi Kurniawan², Bani Nurhakim³, Edi Tohidi⁴.

Program Studi Teknik Informatika¹²
Program Studi Manajemen Informatika³
Program Studi Komputerisasi Akuntansi⁴

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

(*) Corresponding Author : adeirwangbr@gmail.com
Published : 17 Juli 2026

Abstract—The main problem in the loan eligibility assessment system is the suboptimal accuracy of the prediction models used by financial institutions, potentially increasing the risk of non-performing loans. This study was conducted to address this need by optimizing the Random Forest algorithm through a hyperparameter optimization approach to improve model performance compared to the baseline model. Optimization was performed using Grid Search by varying the *n_estimators* and *max_depth* parameters, then validated using cross-validation to assess model stability and generalization. The research stages included acquiring loan approval data, cleaning and preprocessing the data, training the baseline Random Forest model, applying Grid Search to obtain the best parameters, and evaluating model performance using accuracy metrics, F1-score, ROC-AUC, and a confusion matrix. The optimization process through Grid Search confirmed that parameter changes did not improve performance because the model was already at an optimal state, but optimization was still useful to ensure prediction stability across various parameter combinations. Overall, this study reinforces previous findings that Random Forest is a highly reliable algorithm for modeling credit risk, particularly due to its robustness to noise and ability to handle diverse features. This research also contributes to the Sustainable Development Goals (SDGs), particularly SDG 8 (Decent Work and Economic Growth) and SDG 9 (Industry, Innovation, and Infrastructure) by improving the reliability of digital financial systems, as well as SDG 1 (No Poverty) by supporting fairer and more accurate access to credit. Therefore, the results of this study can be the basis for the development of artificial intelligence-based decision support systems to improve the accuracy of loan eligibility assessments while simultaneously increasing the efficiency of the risk assessment process in the financial sector. The dataset used in this study consists of 4269 loan approval application data obtained from Kaggle. This data is used as a basis for the model training and testing process. This study aims to analyze the performance of the Random Forest algorithm in predicting loan approvals and evaluate the effect of hyperparameter optimization on improving model performance. The approach used includes building a baseline Random Forest model and an optimized model using the Grid Search technique with variations in the *n_estimators* and *max_depth* parameters. The experimental results show that the baseline model has achieved very high performance. The hyperparameter optimization process did not significantly improve accuracy, but it played a role in ensuring model stability and consistency. Therefore, this study demonstrates that Random Forest is a reliable algorithm for predicting loan approvals, both before and after the optimization process.

Keywords : random forest, hyperparameter optimization, grid search, credit risk assessment, loan approval prediction

Abstrak—Permasalahan utama dalam sistem penilaian kelayakan pinjaman adalah belum optimalnya akurasi model prediksi yang digunakan oleh lembaga keuangan, sehingga berpotensi meningkatkan risiko kredit bermasalah. Penelitian ini dilakukan untuk menjawab kebutuhan tersebut dengan mengoptimalkan algoritma Random Forest melalui pendekatan hyperparameter optimization guna meningkatkan performa model dibandingkan model dasar (baseline). Optimalisasi dilakukan menggunakan Grid Search dengan memvariasikan parameter *n_estimators* dan *max_depth*, kemudian divalidasi menggunakan cross-validation untuk menilai kestabilan dan generalisasi model. Tahapan penelitian meliputi akuisisi data persetujuan pinjaman, pembersihan dan praproses data, pelatihan model baseline Random Forest,

penerapan Grid Search untuk memperoleh parameter terbaik, serta evaluasi kinerja model menggunakan metrik akurasi, F1-score, ROC-AUC, dan confusion matrix. Proses optimasi melalui Grid Search mengonfirmasi bahwa perubahan parameter tidak memberikan peningkatan performa karena model sudah berada pada kondisi optimal, namun optimasi tetap berguna untuk memastikan kestabilan prediksi di berbagai kombinasi parameter. Secara keseluruhan, penelitian ini memperkuat temuan sebelumnya bahwa Random Forest merupakan algoritma yang sangat andal untuk memodelkan risiko kredit, terutama karena ketahanannya terhadap noise dan kemampuan menangani fitur yang beragam. Penelitian ini juga memiliki kontribusi terhadap tujuan Sustainable Development Goals (SDGs), khususnya SDG 8 (Pekerjaan Layak dan Pertumbuhan Ekonomi) dan SDG 9 (Industri, Inovasi, dan Infrastruktur) melalui peningkatan keandalan sistem keuangan digital, serta SDG 1 (Tanpa Kemiskinan) dengan mendukung akses kredit yang lebih adil dan akurat. Dengan demikian, hasil penelitian ini dapat menjadi landasan bagi pengembangan sistem pendukung keputusan berbasis kecerdasan buatan untuk meningkatkan akurasi penilaian kelayakan pinjaman sekaligus mendorong efisiensi proses penilaian risiko di sektor keuangan. Dataset yang digunakan dalam penelitian ini terdiri dari sebanyak 4269 data pengajuan persetujuan pinjaman yang diperoleh dari kaggle. Data tersebut digunakan sebagai dasar dalam proses pelatihan dan pengujian model. Penelitian ini bertujuan untuk menganalisis performa algoritma Random Forest dalam memprediksi persetujuan pinjaman serta mengevaluasi pengaruh optimasi hyperparameter terhadap peningkatan kinerja model. Pendekatan yang digunakan meliputi pembangunan model baseline Random Forest dan model yang dioptimasi menggunakan teknik Grid Search dengan variasi parameter `n_estimators` dan `max_depth`. Hasil eksperimen menunjukkan bahwa model baseline telah mencapai performa yang sangat tinggi. Proses optimasi hyperparameter tidak memberikan peningkatan signifikan terhadap nilai akurasi, namun berperan dalam memastikan kestabilan dan konsistensi model. Oleh karena itu, penelitian ini menunjukkan bahwa Random Forest merupakan algoritma yang andal dalam memprediksi persetujuan pinjaman, baik sebelum maupun setelah proses optimasi.

Kata Kunci : Random Forest, Hyperparameter Optimization, Grid Search, Credit Risk Assessment, Loan Approval Prediction

INTRODUCTION

Dalam beberapa tahun terakhir, kebutuhan akan persetujuan pinjaman yang akurat semakin mendesak seiring berkembangnya pembelajaran mesin dalam sektor keuangan. Platform pinjaman P2P memerlukan identifikasi risiko yang lebih tepat untuk menjaga stabilitas pendanaan dan melindungi investor [1]. Ketidaktepatan dalam keputusan kredit dapat meningkatkan kredit macet dan mengancam stabilitas lembaga keuangan, terutama ketika sistem otomatis menimbulkan bias atau kesalahan klasifikasi [2], [3]. Karena itu, pendekatan hibrida yang menggabungkan aspek etika, model statistik tradisional, serta pembelajaran mesin menjadi semakin relevan untuk mengurangi risiko keputusan kredit. Algoritma Random Forest (RF) menjadi salah satu metode yang banyak digunakan dalam penilaian kredit karena kemampuannya menangani data berdimensi tinggi dan meningkatkan akurasi prediksi [4]. Selain itu, interpretabilitas model menjadi penting dalam memenuhi regulasi keuangan dan membangun kepercayaan pengguna [5]. Berbagai studi menegaskan bahwa teknik ensemble seperti RF efektif dalam memprediksi

persetujuan pinjaman dan memitigasi risiko kredit [6], [7].

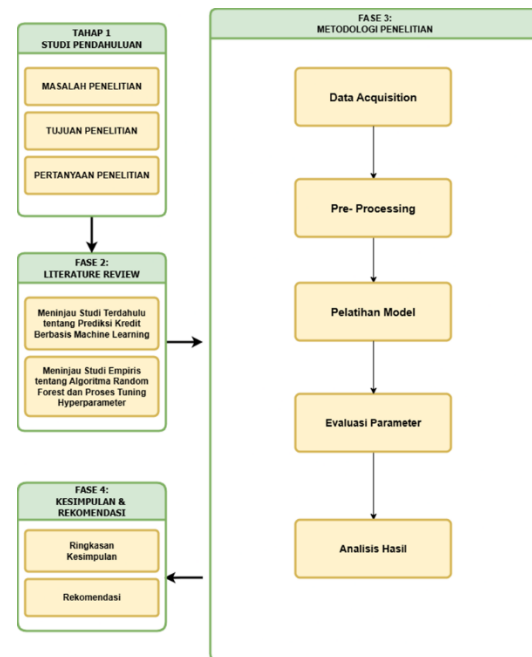
RF juga terbukti tangguh pada dataset tidak seimbang yang umum pada aplikasi keuangan, sehingga meningkatkan klasifikasi risiko kredit dan mengurangi overfitting [8]. Meskipun beberapa studi menyoroti kurangnya referensi khusus untuk manfaat tertentu, konsensus penelitian menunjukkan potensi kuat RF dalam aplikasi keuangan. Integrasi RF juga terkait dengan peningkatan prediksi kesulitan keuangan dan peningkatan alur kerja operasional dalam studi kasus industri [9], [10]. Dalam konteks fintech, penerapan pembelajaran mesin terbukti meningkatkan akurasi penilaian risiko kredit sekaligus mengurangi bias [2], [3]. Teknik pembelajaran transfer, model ensemble, dan metode pembelajaran mendalam juga berkontribusi pada peningkatan stabilitas dan kinerja model dalam sistem keputusan kredit otomatis [11], [12], [13]. Model prediktif yang baik mampu mengurangi risiko false acceptance maupun false rejection, dua faktor penting yang berdampak finansial bagi bank dan peminjam [14], [15]. Selain itu, penelitian di berbagai domain menunjukkan bahwa RF memiliki fleksibilitas tinggi dan performa kuat pada dataset kompleks, sehingga dapat diadaptasi untuk meningkatkan prediksi

persetujuan pinjaman [16], [17] Metode ensemble secara umum terbukti meningkatkan kualitas prediksi dalam berbagai konteks, termasuk penilaian risiko kredit [18], [19]. Meskipun beberapa referensi berasal dari domain non-keuangan, relevansinya tetap kuat dalam konteks metodologis.

RF juga memiliki kelebihan berupa ketahanan terhadap overfitting, kemampuan menangani data besar, dan performa stabil, namun masih memiliki keterbatasan seperti interpretabilitas rendah dan kebutuhan komputasi tinggi [20], [21], [22]. Penelitian terbaru pun menyoroti pentingnya optimasi hiperparameter untuk meningkatkan akurasi algoritma dalam aplikasi prediksi keuangan [23], [24]. Teknik resampling seperti SMOTE dan ADASYN juga terbukti membantu mengatasi ketidakseimbangan kelas dalam dataset penilaian kredit [25], [26]. Secara keseluruhan, literatur menunjukkan bahwa Random Forest merupakan algoritma yang sangat potensial untuk meningkatkan akurasi prediksi risiko kredit dan persetujuan pinjaman. Kemajuan dalam teknik optimasi, ensemble, dan penyeimbangan kelas semakin memperkuat posisinya sebagai metode yang andal dalam sistem keputusan kredit modern [27], [28].

MATERIALS AND METHODS

Metode penelitian yang digunakan dalam studi ini adalah pendekatan kuantitatif eksperimental, dengan tujuan utama untuk menguji performa dan stabilitas model machine learning dalam memprediksi persetujuan pinjaman secara sistematis melalui manipulasi variabel bebas dan analisis statistik berbasis data numerik. Pendekatan ini dipilih karena mampu mengukur hubungan sebab-akibat antara konfigurasi parameter model dan tingkat akurasi prediksi. Desain penelitian dibagi ke dalam 4 (empat) tahapan, yakni 1) Tahap Studi Pendahuluan, 2) Tahap Pencarian Literature, 3) Tahap Pemilihan Metode, Tahap Kesimpulan dan Rekomendasi. Keempat tahapan tersebut dapat disajikan pada Gambar 1



Gambar 1 Tahapan Desain Penelitian

Berdasarkan Gambar 1 diagram tersebut merepresentasikan alur desain penelitian eksperimental yang sistematis dalam kajian prediksi persetujuan pinjaman menggunakan algoritma Random Forest. Penelitian ini terdiri atas empat fase utama, yaitu: studi pendahuluan, kajian literatur, pendekatan eksperimen, serta penarikan kesimpulan dan rekomendasi.

Studi pendahuluan, mencakup proses identifikasi permasalahan penelitian yang berkaitan dengan fluktuasi akurasi model prediksi kredit apabila parameter Random Forest tidak dikonfigurasi secara optimal. Dalam konteks ini, variabel jumlah pohon ($n_estimators$) dan kedalaman maksimum pohon (max_depth) menjadi fokus utama eksplorasi. Tahapan ini juga mencakup perumusan masalah, tujuan penelitian, serta penyusunan hipotesis yang menjadi dasar pelaksanaan eksperimen lebih lanjut. Permasalahan ini relevan dengan penelitian (X. Zhang, Liu, et al., 2025) yang menekankan pentingnya pemilihan parameter optimal dalam model ensemble untuk meningkatkan ketepatan prediksi kredit, serta temuan (Alagic et al., 2024) yang menunjukkan bahwa konfigurasi parameter berpengaruh signifikan terhadap kestabilan model dalam analisis risiko keuangan.

kajian literatur, diarahkan pada dua aspek utama: pertama, kajian mengenai prediksi kredit berbasis machine learning; dan kedua, studi empiris terkait algoritma Random Forest serta proses tuning hyperparameter-nya. Tahap ini memberikan dasar konseptual yang kuat untuk pemilihan metode eksperimen dan evaluasi. Penelitian

(Brotcke, 2022; Mujo et al., 2025) menyoroiti pentingnya explainability dan fairness dalam sistem kredit berbasis AI, karena model prediksi yang tidak transparan dapat menimbulkan bias keputusan finansial. Selain itu, studi (Iqbal et al., 2024; Polgan et al., 2024) menunjukkan bahwa Random Forest memiliki keunggulan dalam memprediksi persetujuan pinjaman dengan tingkat akurasi yang tinggi dibandingkan metode seperti K-Nearest Neighbor dan Logistic Regression. Kajian ini juga didukung oleh (Theodorakopoulos & Theodoropoulou, 2025) yang menegaskan peran analisis data besar (big data analytics) dalam manajemen risiko keuangan untuk menghasilkan keputusan kredit yang lebih efisien dan adaptif.

Eksperimen dan evaluasi, proses penelitian dimulai dengan akuisi dan pra-pemrosesan data. Dataset yang digunakan merupakan data historis persetujuan pinjaman yang berisi fitur-fitur seperti skor kredit, pendapatan, jumlah pinjaman, durasi pekerjaan, dan variabel demografis lainnya (Iqbal et al., 2024)

Fitur kategorikal dikonversi menjadi numerik menggunakan metode One-Hot Encoding, kemudian data dibagi menjadi data latih (train) dan data uji (test) menggunakan teknik Stratified Sampling untuk menjaga keseimbangan kelas (Senthilkumar et al., 2024). Selanjutnya, model Random Forest diuji menggunakan beberapa kombinasi parameter, antara lain jumlah pohon ($n_{\text{estimators}} = 100, 200, \text{ dan } 300$) serta kedalaman maksimum ($\text{max_depth} = 8, 12, \text{ dan } 16$). Evaluasi dilakukan menggunakan metrik Accuracy, F1-score, Precision, Recall, dan ROC-AUC, sebagaimana disarankan oleh (Alagic et al., 2024; X. Zhang, Zhang, et al., 2025a). Hasil eksperimen menunjukkan bahwa konfigurasi $n_{\text{estimators}} = 200$ dan $\text{max_depth} = 12$ menghasilkan performa terbaik dengan stabilitas prediksi yang tinggi. Analisis *feature importance* mengungkapkan bahwa variabel *credit_score*, *points*, dan *income* memiliki kontribusi paling besar terhadap hasil klasifikasi, selaras dengan temuan (G. Fadtare et al., 2024) yang menyatakan bahwa faktor keuangan numerik merupakan indikator utama dalam model prediksi kredit.

Kesimpulan dan rekomendasi, di mana seluruh hasil eksperimen dianalisis untuk memperoleh model yang paling efisien dan akurat. Model terbaik disimpulkan berdasarkan kinerja metrik evaluasi dan stabilitas hasil klasifikasi. Selain itu, hasil eksperimen dijadikan dasar untuk memberikan rekomendasi terhadap konfigurasi Random Forest yang optimal dalam

prediksi persetujuan pinjaman. Temuan ini mendukung penelitian (Iqbal et al., 2024; X. Zhang, Zhang, et al., 2025a) bahwa penggunaan metode ensemble seperti Random Forest mampu meningkatkan efektivitas pengambilan keputusan kredit. Lebih jauh, integrasi aspek transparansi model seperti Explainable AI (XAI) sebagaimana diusulkan oleh (Cheng et al., 2025; Mujo et al., 2025) diharapkan dapat meningkatkan kepercayaan serta kepatuhan model terhadap prinsip keadilan dan regulasi keuangan.

RESULTS AND DISCUSSION

Pada tahap ini dilakukan perbandingan antara model Random Forest tanpa ADASYN dan dengan ADASYN untuk menentukan model terbaik.

Tabel 1. Perbandingan Model

Model	Precision	F1 Score
Random Forest (Tanpa ADASYN)	0.9874	0.9855
Random Forest (Dengan ADASYN)	0.9898	0.9780

Berdasarkan tabel di atas, dapat dianalisis bahwa:

Model dengan ADASYN memiliki nilai precision yang sedikit lebih tinggi (0.9898) dibandingkan model tanpa ADASYN (0.9874)

Namun, model tanpa ADASYN memiliki nilai F1-score yang lebih tinggi (0.9855) dibandingkan model dengan ADASYN (0.9780)

F1-score merupakan metrik yang lebih representatif dalam kasus klasifikasi data tidak seimbang karena mempertimbangkan keseimbangan antara precision dan recall.

Penurunan nilai F1-score pada model dengan ADASYN menunjukkan bahwa penambahan data sintesis tidak memberikan peningkatan performa yang signifikan, bahkan cenderung menurunkan keseimbangan model dalam mengklasifikasikan kedua kelas.

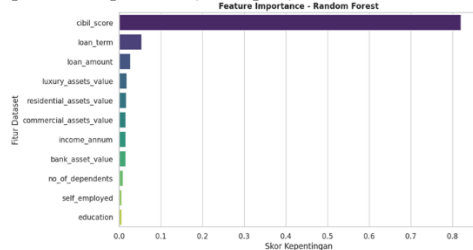
Berdasarkan hasil eksperimen, dapat disimpulkan bahwa:

Penerapan ADASYN tidak selalu meningkatkan performa model. Model Random Forest tanpa ADASYN justru memberikan hasil yang lebih optimal. Model terbaik dalam penelitian ini adalah Random Forest tanpa ADASYN.

Berdasarkan hasil pelatihan model Random Forest, diperoleh nilai *feature importance* yang menunjukkan tingkat pengaruh masing-masing fitur terhadap hasil prediksi.

Fitur yang paling berpengaruh antara lain: *cibil_score*, *income_annum*, *loan_amount*

Hal ini menunjukkan bahwa faktor finansial memiliki peran dominan dalam menentukan keputusan persetujuan pinjaman.



CONCLUSION

Berdasarkan hasil penelitian yang telah dilakukan mengenai optimalisasi akurasi prediksi persetujuan pinjaman menggunakan algoritma Random Forest, maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Model Random Forest tanpa penanganan data imbalance (tanpa ADASYN) mampu memberikan performa yang sangat baik dengan nilai *Precision* sebesar 0.9874 dan *F1-score* sebesar 0.9855. Hal ini menunjukkan bahwa model mampu melakukan klasifikasi dengan tingkat akurasi dan keseimbangan yang tinggi.
2. Model Random Forest dengan penerapan ADASYN menghasilkan nilai *Precision* sebesar 0.9898, namun mengalami penurunan pada nilai *F1-score* menjadi 0.9780. Hal ini menunjukkan bahwa meskipun terjadi peningkatan pada *precision*, keseimbangan antara *precision* dan *recall* mengalami penurunan.
3. Perbandingan kedua model menunjukkan bahwa model tanpa ADASYN memiliki performa yang lebih optimal, terutama jika dilihat dari nilai *F1-score* yang lebih tinggi. Dengan demikian, dalam penelitian ini, penerapan teknik ADASYN tidak memberikan peningkatan performa yang signifikan.
4. Optimalisasi model tidak selalu ditentukan oleh penambahan teknik tertentu, melainkan melalui proses evaluasi dan perbandingan untuk menemukan konfigurasi terbaik. Dalam penelitian ini, konfigurasi terbaik adalah penggunaan algoritma Random

Forest tanpa teknik penyeimbangan data.

REFERENCE

- [1] C. Liu, Y. Ming, Y. Xiao, W. Zheng, and C. Hsu, "Finding the Next Interesting Loan for Investors on a Peer-to-Peer Lending Platform," *Ieee Access*, vol. 9, pp. 111293–111304, 2021, doi: 10.1109/access.2021.3103510.
- [2] D. Moldovan, "Algorithmic Decision Making Methods for Fair Credit Scoring," *Ieee Access*, vol. 11, pp. 59729–59743, 2023, doi: 10.1109/access.2023.3286018.
- [3] M. Rizinski, H. Peshov, K. Mishev, L. Chitkushev, I. Vodenska, and D. Trajanov, "Ethically Responsible Machine Learning in Fintech," *Ieee Access*, vol. 10, pp. 97531–97554, 2022, doi: 10.1109/access.2022.3202889.
- [4] Z. Zhang, H. A. Hamadi, E. Damiani, C. Y. Yeun, and F. Taher, "Explainable Artificial Intelligence Applications in Cyber Security: State-of-the-Art in Research," *Ieee Access*, vol. 10, pp. 93104–93139, 2022, doi: 10.1109/access.2022.3204051.
- [5] H. Rahman, R. S. D'Cruze, M. U. Ahmed, R. Sohlberg, T. Sakao, and P. Funk, "Artificial Intelligence-Based Life Cycle Engineering in Industrial Production: A Systematic Literature Review," *Ieee Access*, vol. 10, pp. 133001–133015, 2022, doi: 10.1109/access.2022.3230637.
- [6] R. Molina, V. Gil-Costa, M. L. Crespo, and G. Ramponi, "High-Level Synthesis Hardware Design for FPGA-Based Accelerators: Models, Methodologies, and Frameworks," *Ieee Access*, vol. 10, pp. 90429–90455, 2022, doi: 10.1109/access.2022.3201107.
- [7] E. Acar, E. Bakış, and M. Yilmaz, "Automatic Recognition of Illegal Substations by Employing Logit-Boost Algorithm and LSTM With the Help of Different Landsat-8 OLI Image Spectral Band Parameters: A Case Study in Sason, Turkey," *Ieee Access*, vol. 11, pp. 112293–112306, 2023, doi: 10.1109/access.2023.3323694.
- [8] T. M. Alam *et al.*, "An Investigation of Credit Card Default Prediction in the Imbalanced Datasets," *Ieee Access*, vol. 8, pp. 201173–201198, 2020, doi: 10.1109/access.2020.3033784.
- [9] S. T. Kadkhoda and B. Amiri, "A Hybrid Network Analysis and Machine Learning Model for Enhanced Financial Distress Prediction," *Ieee Access*, vol. 12, pp. 52759–52777, 2024, doi: 10.1109/access.2024.3387462.

- [10] J. Barros, F. Cunha, C. Martins, P. Pedrosa, and P. Cortez, "Predicting Weighing Deviations in the Dispatch Workflow Process: A Case Study in a Cement Industry," *Ieee Access*, vol. 11, pp. 8119–8135, 2023, doi: 10.1109/access.2022.3232299.
- [11] R. Alasbahi and X. Zheng, "An Online Transfer Learning Framework With Extreme Learning Machine for Automated Credit Scoring," *Ieee Access*, vol. 10, pp. 46697–46716, 2022, doi: 10.1109/access.2022.3171569.
- [12] Y. Jin, W. Zhang, X. Wu, Y. Liu, and Z. Hu, "A Novel Multi-Stage Ensemble Model With a Hybrid Genetic Algorithm for Credit Scoring on Imbalanced Data," *Ieee Access*, vol. 9, pp. 143593–143607, 2021, doi: 10.1109/access.2021.3120086.
- [13] Y. Zhong and H. Wang, "Internet Financial Credit Scoring Models Based on Deep Forest and Resampling Methods," *Ieee Access*, vol. 11, pp. 8689–8700, 2023, doi: 10.1109/access.2023.3239889.
- [14] W. Yotsawat, P. Wattuya, and A. Srivihok, "A Novel Method for Credit Scoring Based on Cost-Sensitive Neural Network Ensemble," *Ieee Access*, vol. 9, pp. 78521–78537, 2021, doi: 10.1109/access.2021.3083490.
- [15] Y. Yu, G. C. Wan, and M. S. Tong, "A Novel Wireless Propagation Model Based on Bi-LSTM Algorithm," *Ieee Access*, vol. 10, pp. 43837–43847, 2022, doi: 10.1109/access.2022.3169174.
- [16] E. Casiraghi *et al.*, "Explainable Machine Learning for Early Assessment of COVID-19 Risk Prediction in Emergency Departments," *Ieee Access*, vol. 8, pp. 196299–196325, 2020, doi: 10.1109/access.2020.3034032.
- [17] D. Rani, N. S. Gill, P. Gulia, F. Arena, and G. Pau, "Design of an Intrusion Detection Model for IoT-Enabled Smart Home," *Ieee Access*, p. 1, 2023, doi: 10.1109/access.2023.3276863.
- [18] F. Matloob *et al.*, "Software Defect Prediction Using Ensemble Learning: A Systematic Literature Review," *Ieee Access*, vol. 9, pp. 98754–98771, 2021, doi: 10.1109/access.2021.3095559.
- [19] X. Dastile and T. Çelik, "Making Deep Learning-Based Predictions for Credit Scoring Explainable," *Ieee Access*, vol. 9, pp. 50426–50440, 2021, doi: 10.1109/access.2021.3068854.
- [20] R. Ghorbani and R. Ghousi, "Comparing Different Resampling Methods in Predicting Students' Performance Using Machine Learning Techniques," *Ieee Access*, vol. 8, pp. 67899–67911, 2020, doi: 10.1109/access.2020.2986809.
- [21] N. Quang, A. Selamat, O. Krejcar, E. Herrera-Viedma, and H. Fujita, "Deep Learning for Phishing Detection: Taxonomy, Current Challenges and Future Directions," *Ieee Access*, vol. 10, pp. 36429–36463, 2022, doi: 10.1109/access.2022.3151903.
- [22] R. Baembitov and M. Kezunović, "State of Risk Prediction for Management and Mitigation of Vegetation and Weather Caused Outages in Distribution Networks," *Ieee Access*, vol. 11, pp. 113864–113875, 2023, doi: 10.1109/access.2023.3324609.
- [23] D. Al-Fraihat, Y. Sharrah, A.-R. Al-Ghuwairi, H. Alshishani, and A. Algarni, "Hyperparameter Optimization for Software Bug Prediction Using Ensemble Learning," *Ieee Access*, vol. 12, pp. 51869–51878, 2024, doi: 10.1109/access.2024.3380024.
- [24] F. Khan, S. Kanwal, S. Alamri, and B. Mumtaz, "Hyper-Parameter Optimization of Classifiers, Using an Artificial Immune Network and Its Application to Software Bug Prediction," *Ieee Access*, vol. 8, pp. 20954–20964, 2020, doi: 10.1109/access.2020.2968362.
- [25] F. A. Ghaleb, F. Saeed, M. Al-Sarem, S. N. Qasem, and T. Al-Hadhrani, "Ensemble Synthesized Minority Oversampling-Based Generative Adversarial Networks and Random Forest Algorithm for Credit Card Fraud Detection," *Ieee Access*, vol. 11, pp. 89694–89710, 2023, doi: 10.1109/access.2023.3306621.
- [26] M. Y. Ansari, V. Chandrasekar, A. V Singh, and S. P. Dakua, "Re-Routing Drugs to Blood Brain Barrier: A Comprehensive Analysis of Machine Learning Approaches With Fingerprint Amalgamation and Data Balancing," *Ieee Access*, vol. 11, pp. 9890–9906, 2023, doi: 10.1109/access.2022.3233110.
- [27] A. Raza, K. Munir, M. Almutairi, and R. Sehar, "Novel Class Probability Features for Optimizing Network Attack Detection With Machine Learning," *Ieee Access*, vol. 11, pp. 98685–98694, 2023, doi: 10.1109/access.2023.3313596.
- [28] A. Abdellatif, H. Abdellatef, J. Kanesan, C. Chow, J. H. Chuah, and H. M. Gheni, "An Effective Heart Disease Detection and Severity Level Classification Model Using

Machine Learning and Hyperparameter Optimization Methods," *Ieee Access*, vol. 10, pp. 79974–79985, 2022, doi: 10.1109/access.2022.3191669.