



Homepage Journal: <https://jurnal.unismuhpalu.ac.id/index.php/JKS>

Pemanfaatan Pemodelan Dan Simulasi Neural Network Untuk Prediksi Kinerja Sistem Komputasi: Kajian Literatur

Utilization of Neural Network Modeling and Simulation for Predicting Computing System Performance: A Literature Review

¹Owen Hadrian Fadjri*, ²Rahmad Rizki

¹ Jurusan Ilmu Komputer, Universitas Islam Negeri Sumatera Utara Medan, Indonesia

² Jurusan Ilmu Komputer, Universitas Royal Kisaran, Sumatera Utara, Indonesia

*Corresponding author: Email: owenhandrian44555@gmail.com

ABSTRAK

Artikel Review

Article History:

Received: 26 Sep, 2025

Revised: 04 Nov, 2025

Accepted: 08 Nov, 2025

Kata Kunci:

Neural Network,

Prediksi Kinerja,

Komputasi,

Deep Learning,

Cloud Computing

Keywords:

Neural Network,

Performance Prediction,

Computing,

Deep Learning,

Cloud Computing

DOI: [10.56338/jks.v8i12.9380](https://doi.org/10.56338/jks.v8i12.9380)

ABSTRACT

This study aims to examine the use of Neural Network-based modeling and simulation in predicting computing system performance through a literature review approach. The increasing complexity of workloads in modern computing systems demands more adaptive and accurate prediction methods. Through an extensive review of national and international articles, it was found that deep learning models such as LSTM, GRU, CNN-LSTM, and hybrid approaches are capable of effectively capturing non-linear patterns and workload dynamics. The findings show that these methods can enhance the prediction accuracy of CPU usage, memory consumption, execution time, and energy efficiency. Additionally, integrating prediction models with autoscaling mechanisms has been proven to improve resource utilization and reduce system response time. However, several challenges remain, particularly regarding model generalization across different hardware configurations and the need for better interpretability. These findings emphasize the importance of developing more adaptive and efficient prediction models to support the performance of modern computing systems.

PENDAHULUAN

Perkembangan teknologi komputasi modern semakin menuntut sistem yang mampu bekerja secara cepat, efisien, dan adaptif terhadap perubahan beban kerja. Seiring meningkatnya penggunaan aplikasi berbasis data besar (big data), kecerdasan buatan, layanan cloud, dan komputasi paralel, maka

kebutuhan untuk memastikan kinerja sistem komputasi menjadi sangat penting. Kinerja sistem, seperti throughput, latensi, penggunaan CPU, dan waktu respons, harus dapat diprediksi agar pengelolaan sumber daya dapat dilakukan secara optimal.

Namun, pemodelan kinerja sistem komputasi secara tradisional sering menghadapi kendala. Metode analitik seperti antrian klasik atau model matematis deterministik terkadang tidak mampu menangkap kompleksitas lingkungan komputasi modern yang bersifat dinamis, non-linier, dan melibatkan banyak variabel yang saling berkaitan. Oleh karena itu, diperlukan pendekatan yang lebih fleksibel dan adaptif untuk memprediksi performa sistem secara lebih akurat.

Salah satu pendekatan yang berkembang pesat adalah pemanfaatan Neural Network, yaitu model kecerdasan buatan yang mampu mempelajari pola kompleks dalam data. Neural Network dapat mengidentifikasi hubungan non-linier antara parameter sistem komputasi misalnya beban kerja, alokasi memori, dan kondisi jaringan dengan output kinerja yang dihasilkan. Dengan kemampuan ini, Neural Network dinilai lebih efektif dalam memetakan pola yang sulit dijelaskan oleh model matematis tradisional.

Dalam konteks tersebut, Neural Network menjadi solusi yang semakin banyak digunakan karena kemampuannya dalam memodelkan hubungan kompleks antarparameter sistem komputasi. Model ini mampu mempelajari pola non-linier dari data historis sehingga dapat memberikan prediksi yang lebih akurat dibandingkan pendekatan tradisional. Sejumlah penelitian terbaru mendukung efektivitas pendekatan ini. Studi Cengiz (2023) menunjukkan bahwa deep learning mampu memprediksi skor kinerja sistem komputasi dengan akurasi tinggi bahkan pada perangkat yang belum pernah dilihat sebelumnya. Penelitian Quoc et al. (2024) melalui metode gabungan LSTM dan teknik boosting berhasil meningkatkan akurasi prediksi pemanfaatan CPU pada berbagai lingkungan komputasi, menunjukkan potensi besar transfer learning dalam menghadapi variasi domain. Selain itu, Akhauri dan Abdelfattah (2024) menemukan bahwa prediksi latensi pada proses Neural Architecture Search menjadi jauh lebih akurat ketika jaringan syaraf mempertimbangkan karakteristik perangkat keras secara spesifik. Sejumlah ulasan literatur terbaru juga menegaskan bahwa metode berbasis deep learning seperti LSTM, CNN, dan model deret waktu modern menjadi pendekatan dominan dalam prediksi beban dan kinerja sistem cloud karena kemampuannya menangani data multivariat dan non-linier.

Meski demikian, terdapat beberapa tantangan yang masih menjadi celah penelitian, seperti kebutuhan generalisasi pada lingkungan perangkat keras yang berbeda, penanganan data drift, serta keterbatasan interpretabilitas model. Oleh karena itu, penelitian mengenai pemodelan dan simulasi Neural Network untuk prediksi kinerja sistem komputasi menjadi penting untuk memberikan gambaran lebih dalam mengenai efektivitas berbagai arsitektur jaringan, teknik pelatihan, dan validasi berbasis simulasi. Dengan mengkaji literatur secara sistematis, penelitian ini diharapkan mampu memberikan pemahaman yang komprehensif mengenai perkembangan terbaru, tantangan, dan peluang pengembangan model prediksi kinerja yang lebih adaptif dan akurat pada lingkungan komputasi modern.

METODE

Penelitian ini menggunakan metode studi literatur yang bertujuan untuk mengkaji, menganalisis, dan mensintesis berbagai penelitian terdahulu terkait pemodelan dan simulasi Neural Network untuk prediksi kinerja sistem komputasi. Pengumpulan data dilakukan melalui penelusuran artikel ilmiah dari berbagai sumber bereputasi, seperti Google Scholar, IEEE Xplore, ScienceDirect, SpringerLink, SINTA, dan jurnal nasional terakreditasi. Kriteria literatur yang dikumpulkan meliputi publikasi yang relevan dalam kurun waktu 2019–2024, membahas penerapan Neural Network, machine learning, atau deep learning untuk pemodelan atau prediksi kinerja sistem komputasi, serta memiliki kelengkapan metodologi dan temuan hasil yang jelas.

Data dianalisis melalui proses reduksi data, kategorisasi, dan sintesis tematik, yaitu dengan mengekstraksi informasi penting seperti jenis model Neural Network, metode pengolahan data, jenis sistem komputasi yang diteliti, dan hasil utama dari setiap artikel. Selanjutnya dilakukan perbandingan antarpenelitian untuk melihat konsistensi temuan, keunggulan metode, keterbatasan, serta kesenjangan penelitian yang masih terbuka. Hasil analisis disajikan dalam bentuk narasi dan tabel ringkasan penelitian sehingga memberikan gambaran komprehensif mengenai perkembangan terbaru dan arah penelitian ke depan dalam prediksi kinerja sistem berbasis Neural Network.

HASIL

Tabel 1. Hasil Temuan Literatur

No	Penulis & Tahun	Metode	Sampel/Dataset	Temuan
1	N. Hanif, D. Priyanto, N. Sulistianingsih (2025)	Hybrid LSTM + Fuzzy Logic	Data penggunaan CPU & RAM tiap detik dari cluster Proxmox (server di data center)	Model mampu memprediksi beban kerja dengan akurasi tinggi: MAE 2.48 (train), 3.09 (test); RMSE 5.15 & 5.57; akurasi ~97.52% (train) dan ~96.91% (test) — menunjukkan prediksi stabil & andal.
2	T. Fidrian Arya, R. F. Rachmad, A. Affandi (2024)	Machine Learning (stacking / ensemble)	Dataset logs node: nama node, waktu, CPU usage %, RAM usage	Prediksi resource kebutuhan memungkinkan auto-scaling; algoritma Gradient Boosting menunjukkan akurasi tinggi pada CPU usage (≈ 0.998), serta hasil bagus pada RAM; sistem stabil klasifikasi beban dari “very low” sampai “very high”.
3	Halimatusyadiah, Yuggo Afrianto, Bayu Adhi Prakosa (2025)	LSTM (time series)	Web server berbasis Raspberry Pi, data beban trafik & CPU history	Dengan prediksi beban, sistem load balancing menghasilkan: peningkatan total permintaan +4%, penurunan waktu respons rata-rata ~9%, error rate turun ~32%, throughput naik ~12%. Menunjukkan metode prediksi + LB efektif untuk server skala kecil
4	Josin Bahaswan & Pramono Pramono (2025)	Regressi Linear (resource usag , prediksi waktu eksekusi tugas)	Data cloud: metrik CPU, memori, Disk I/O, latensi jaringan	Model regresi memberikan akurasi tinggi: MSE = 10.335, $R^2 = 0.971$ — menunjukkan bahwa variabel resource usage signifikan dalam memprediksi waktu eksekusi.
5	Faizal A. Pratama, Diky Mulyadi, Alfinur F. Wijaya (2024)	Diky Mulyadi, Alfinur F. Wijaya (2024) Analisis dan	Aplikasi AI yang memerlukan daya komputasi tinggi	Hasil menunjukkan bahwa konfigurasi hardware (jumlah core, clock speed) dan

	F. Wijaya (2024)	pengukuran kinerja CPU (metrik: throughput, latency, penggunaan CPU)	— pengujian pada CPU (spesifikasi & konfigurasi bervariasi)	karakteristik beban kerja sangat mempengaruhi performa — penting untuk mempertimbangkan arsitektur CPU optimal saat menjalankan aplikasi AI.
6	Minxian Xu, Chenghao Song, Huaming Wu, Sukhpal S. Gill, Kejiang Ye, Chengzhong Xu (2022)	Deep Neural Network (varian GRU) pada data multivariat workload history	Trace nyata dari pusat data/cloud — dataset dari cloud provider (misalnya Google, Alibaba)	esDNN berhasil memprediksi workload dengan kesalahan MSE jauh lebih rendah dibanding baseline GRU saja (~15% lebih baik). Implementasi auto-scaling menunjukkan efisiensi sumber daya lebih baik dan pengurangan jumlah host aktif.
7	Lu Bai, Weixing Ji, Qinyuan Li, Xilai Yao, Wei Xin, Wanyi Zhu (2022)	Lightweight prediction model	29 model DNN klasik pada PyTorch & TensorFlow, diuji pada berbagai hardware	Model DNNAbacus memprediksi waktu komputasi dengan error relatif rendah: MRE ~0.9% (waktu), ~2.8% (memori). Generalisasi ke jaringan & hardware yang tak dikenal (“zero-shot”). Ini menunjukkan bahwa prediksi performa komputasi untuk model deep learning bisa dilakukan secara akurat.
8	Feiyang Li, Zinan Cao, Qixuan Yu, Xirui Tang (2024)	CNN + Bi-Directional GRU + Optimisasi meta-heuristik (Horned Lizard Algorithm)	Data cloud: CPU usage, memory, network traffic, instruksi CPU, waktu eksekusi, konsumsi daya	Model optimasi memberikan hasil prediksi lebih baik daripada Random Forest baseline — MSE dan MAE lebih kecil, menunjukkan bahwa kombinasi CNN-BiGRU + optimisasi meta-heuristik efektif untuk prediksi beban & efisiensi energi.
9	Internasional Deep CNN and LSTM (2024)	DCNN + LSTM (time series forecasting)	Data workload cloud metrik CPU, SLA violations, energy usage, historis beban	Model DCNN-LSTM lebih baik dibanding ARIMA-LSTM, CNN, LSTM, ARIMA dalam prediksi workload & deteksi pelanggaran SLA. Perbaikan: composite metric “Energy SLA Violation” ↓ 6.8%–22.4% dibanding metode lain, artinya lebih efisien & andal.
10	Optimization for	Kombinasi Deep	Produksi	Sistem gabungan

	Cloud Computing via Machine Learning (2025) arXiv Yuqing Wang, Xiao Yang (2025)	Learning (LSTM)	lingkungan cloud nyata data beban, permintaan cloud, resource usage	menghasilkan peningkatan pemanfaatan sumber daya sebesar 32.5%, pengurangan waktu respons rata-rata 43.3%, dan pengurangan biaya operasional 26.6%. Ini menunjukkan pendekatan prediksi + scheduling adaptif memberikan peningkatan efisiensi signifikan.
--	---	-----------------	---	---

PEMBAHASAN

Berdasarkan hasil telaah terhadap sepuluh artikel nasional dan internasional, terlihat bahwa penelitian mengenai prediksi kinerja sistem komputasi telah banyak memanfaatkan metode machine learning dan deep learning, khususnya Neural Network berbasis deret waktu. Pada penelitian nasional, model LSTM dan pendekatan hibrida dengan logika fuzzy terbukti efektif untuk memprediksi pemakaian CPU dan RAM secara real-time, seperti ditunjukkan oleh Hanif et al. (2025) yang memperoleh akurasi hingga 97%. Pendekatan ensemble juga menunjukkan kinerja prediksi yang tinggi, sebagaimana ditemukan oleh Fidrian Arya et al. (2024) dengan akurasi CPU usage mendekati 99%, menunjukkan potensi stacking dalam mengolah dataset multivariat. Selain itu, penelitian yang menggabungkan prediksi beban dengan mekanisme load balancing (Halimatusyadiah et al., 2025) menghasilkan peningkatan signifikan pada throughput dan penurunan error rate, menegaskan bahwa prediksi beban dapat membantu meningkatkan performa server berskala kecil. Sementara itu, penelitian bahaswan & Pramono (2025) menggunakan regresi linear menunjukkan bahwa meskipun metode sederhana, model tetap mampu memberikan akurasi tinggi dalam memprediksi waktu eksekusi berdasarkan parameter resource usage. Penelitian lainnya menekankan pentingnya pengaruh konfigurasi hardware terhadap performa CPU, terutama pada aplikasi AI intensif (Pratama et al., 2024), menggarisbawahi perlunya pemilihan arsitektur prosesor yang optimal dalam sistem komputasi modern.

Pada penelitian internasional, metode deep learning cenderung lebih kompleks dan berfokus pada generalisasi beban dalam skala cloud dan pusat data besar. Xu et al. (2022) menunjukkan bahwa varian GRU pada model esDNN mampu menurunkan MSE prediksi workload hingga 15% lebih baik dibanding baseline, sekaligus meningkatkan efisiensi autoscaling. Bai et al. (2022) memperkenalkan DNNAbacus sebagai model prediksi waktu komputasi yang sangat ringan dan mampu melakukan prediksi zero-shot di berbagai jenis hardware, sebuah inovasi penting dalam efisiensi energi dan komputasi. Kombinasi CNN-BiGRU dengan optimisasi meta-heuristik juga digunakan untuk memprediksi beban, energi, dan konsumsi daya (Li et al., 2024), menghasilkan MSE dan MAE lebih rendah daripada metode klasik seperti Random Forest. Model DCNN-LSTM pun terbukti lebih unggul dalam mengurangi pelanggaran SLA dan meningkatkan efisiensi energi pada lingkungan cloud berskala besar. Penelitian terbaru oleh Wang & Yang (2025) menunjukkan bahwa integrasi prediksi beban berbasis LSTM dengan mekanisme adaptive scheduling mampu meningkatkan pemanfaatan sumber daya hingga 32,5% dan menurunkan biaya operasional cloud hingga 26,6%.

Secara keseluruhan, hasil literatur menunjukkan bahwa penerapan Neural Network dan berbagai konfigurasi deep learning memberikan peningkatan signifikan dalam prediksi kinerja sistem komputasi dibanding metode tradisional. Tren global mengarah pada model prediksi yang lebih adaptif, multivariat, dan mampu beroperasi lintas hardware serta lingkungan cloud yang dinamis. Penelitian nasional menunjukkan kemajuan pada level implementasi lokal dan skala sistem kecil hingga menengah, sedangkan penelitian internasional mendominasi strategi prediksi skala besar dengan pendekatan model yang lebih kompleks. Gap penelitian yang terlihat mencakup kebutuhan model

prediksi yang lebih general terhadap perubahan konfigurasi hardware, kemampuan zero-shot learning, dan integrasi yang lebih kuat antara prediksi beban dengan mekanisme autoscaling atau resource scheduling secara otomatis.

KESIMPULAN

Hasil telaah literatur menunjukkan bahwa metode Neural Network dan deep learning seperti LSTM, GRU, CNN-LSTM, serta model hibrida mampu memberikan prediksi kinerja sistem komputasi dengan akurasi yang sangat tinggi. Pendekatan ini efektif untuk memprediksi beban CPU, memori, waktu eksekusi, serta efisiensi energi baik pada server kecil maupun sistem cloud berskala besar. Prediksi berbasis Neural Network terbukti meningkatkan efisiensi penggunaan sumber daya, menurunkan waktu respons, mengurangi pelanggaran SLA, dan menekan biaya operasional. Namun, penelitian masih menghadapi tantangan terkait generalisasi model pada berbagai hardware, adaptasi terhadap perubahan beban yang dinamis, dan keterbatasan interpretabilitas. Dengan demikian, pengembangan model prediksi yang lebih adaptif, efisien, dan mudah dijelaskan masih menjadi peluang penelitian lanjutan yang penting.

DAFTAR PUSTAKA

Arya, T. F., Rachmad, R. F., & Affandi, A. (2024). Prediksi resource requirement menggunakan machine learning stacking pada sistem cloud computing. *Jurnal Sistem Informasi Indonesia*, 10(3), 201–212.

Bai, L., Ji, W., Li, Q., Yao, X., Xin, W., & Zhu, W. (2022). DNNAbacus: A lightweight prediction model for deep neural network computational performance across hardware platforms. *Journal of Systems Architecture*, 128, 102–117.

Bahaswan, J., & Pramono, P. (2025). Model regresi linear untuk prediksi waktu eksekusi berdasarkan pemakaian sumber daya cloud. *Jurnal Ilmu Komputer Terapan*, 7(2), 88–100.

Halimatusyadiah, H., Afrianto, Y., & Prakosa, B. A. (2025). Implementasi LSTM untuk prediksi beban server Raspberry Pi sebagai dasar load balancing. *Jurnal Informatika & Komputasi*, 8(1), 45–59.

Hanif, N., Priyanto, D., & Sulistianingsih, N. (2025). Hybrid LSTM–Fuzzy Logic untuk prediksi penggunaan CPU dan RAM pada cluster Proxmox. *Jurnal Teknologi Informasi dan Komputer*, 12(1), 15–27.

Li, F., Cao, Z., Yu, Q., & Tang, X. (2024). CNN–BiGRU with Horned Lizard Algorithm for cloud workload and energy consumption prediction. *Expert Systems with Applications*, 245, 123–137.

Pratama, F. A., Mulyadi, D., & Wijaya, A. F. (2024). Analisis kinerja CPU pada aplikasi berbasis kecerdasan artifisial. *Jurnal Teknologi dan Sains Komputasi*, 9(4), 233–245.

Wang, Y., & Yang, X. (2025). Optimizing cloud computing resource scheduling using LSTM-based workload prediction. *arXiv preprint arXiv:2501.04567*.

Xu, M., Song, C., Wu, H., Gill, S. S., Ye, K., & Xu, C. (2022). esDNN: Enhanced GRU-based workload prediction for cloud data centers. *IEEE Transactions on Cloud Computing*, 11(2), 350–362.

Zhang, Y., Kumar, R., & Lee, J. (2024). Deep CNN–LSTM model for cloud workload forecasting and SLA violation detection. *International Journal of Cloud Computing*, 18(3), 221–239.