

Klik disini untuk menuliskan kategori naskah

Optimalisasi Alokasi Sumber Daya pada Arsitektur Private Cloud Menggunakan Algoritma Machine Learning

Andi Zulherry¹, Indah Purnama Sari², Mhd. Basri³, Yohanni Syahra⁴

¹ Fakultas Ilmu Komputer dan Teknologi Informasi, Sains Data, Universitas Muhammadiyah Sumatera Utara, Medan, Indonesia

^{2,3,4} Fakultas Ilmu Komputer dan Teknologi Informasi, Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara, Medan, Indonesia

INFORMASI ARTIKEL

Diterima Redaksi: 00 Januari 00
Revisi Akhir: 00 Februari 00
Diterbitkan Online: 00 Maret 00

KATA KUNCI

private cloud; alokasi sumber daya; Random Forest Regression; machine learning; virtual machine

KORESPONDENSI

Phone: 082273147929
E-mail: andizulherry@umsu.ac.id

A B S T R A K

Private cloud computing telah menjadi solusi infrastruktur yang banyak diadopsi oleh organisasi untuk mengelola sumber daya komputasi secara mandiri dan aman. Namun, pengelolaan alokasi sumber daya seperti CPU, memori, disk, dan bandwidth jaringan pada lingkungan private cloud sering menghadapi tantangan berupa ketidakseimbangan beban kerja, yang dapat menyebabkan kondisi over-provisioning maupun under-provisioning pada virtual machine (VM). Penelitian ini mengusulkan pendekatan optimalisasi alokasi sumber daya berbasis algoritma machine learning untuk memprediksi kebutuhan resource secara dinamis berdasarkan pola penggunaan historis. Metode yang diterapkan meliputi pengumpulan data utilisasi resource dari dataset publik, pra-proses data, pelatihan model Random Forest Regression, serta evaluasi performa model menggunakan metrik Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan koefisien determinasi (R^2). Hasil eksperimen menunjukkan bahwa model mampu memprediksi Disk Write Throughput dengan rata-rata R^2 sebesar 0,81 dan Network Transmitted dengan rata-rata R^2 sebesar 0,46 lintas tiga VM yang diuji. Penelitian ini diharapkan dapat memberikan kontribusi praktis bagi pengelola infrastruktur private cloud dalam merancang sistem alokasi resource yang lebih adaptif dan efisien.

PENDAHULUAN

Pertumbuhan transformasi digital yang pesat di berbagai sektor telah mendorong organisasi untuk mengadopsi cloud computing sebagai infrastruktur utama dalam menjalankan aplikasi dan mengelola data. Di antara model deployment cloud yang tersedia, private cloud menjadi pilihan yang banyak digunakan oleh organisasi yang membutuhkan kontrol lebih besar terhadap keamanan data, kepatuhan regulasi, dan kustomisasi infrastruktur dibandingkan dengan solusi public cloud[1]. Lingkungan private cloud umumnya terdiri dari beberapa virtual machine (VM) yang berjalan di atas server fisik yang dikelola melalui platform virtualisasi, sehingga sumber daya komputasi seperti CPU, memori, disk, dan bandwidth jaringan dapat dibagi dan dialokasikan secara dinamis ke berbagai beban kerja[2].

Meskipun memiliki banyak keunggulan, infrastruktur private cloud masih menghadapi tantangan yang persisten dalam memastikan alokasi sumber daya antar VM berjalan secara efisien. Kebutuhan beban kerja pada suatu VM jarang bersifat statis; kebutuhan tersebut berfluktuasi seiring waktu tergantung pada pola penggunaan aplikasi, aktivitas pengguna, dan siklus bisnis[3]. Ketika alokasi sumber daya dilakukan secara manual atau berdasarkan ambang batas (threshold) yang tetap, hal ini sering menimbulkan dua kondisi yang tidak diinginkan: over-provisioning, yaitu kondisi ketika sumber daya yang dialokasikan melebihi kebutuhan aktual sehingga menjadi idle dan menyebabkan pemborosan kapasitas infrastruktur; serta under-provisioning, yaitu kondisi ketika sumber daya yang dialokasikan tidak mencukupi kebutuhan, sehingga menyebabkan penurunan performa aplikasi dan potensi gangguan layanan[2].

Pendekatan manajemen sumber daya konvensional, seperti alokasi berbasis ambang batas statis atau berbasis aturan (rule-based), umumnya bersifat reaktif daripada prediktif. Pendekatan tersebut menyesuaikan alokasi sumber daya hanya setelah perubahan beban kerja terjadi, bukan mengantisipasi kebutuhan di masa mendatang. Seiring beban kerja pada lingkungan private cloud modern menjadi semakin dinamis dan tidak menentu, pendekatan reaktif semacam itu tidak lagi memadai untuk menjaga efisiensi sumber daya dan kualitas layanan secara optimal[2].

Machine learning menawarkan alternatif yang menjanjikan untuk mengatasi keterbatasan tersebut. Dengan mempelajari pola dari data utilisasi sumber daya historis, algoritma machine learning dapat menghasilkan model prediktif yang mampu memperkirakan kebutuhan sumber daya di masa mendatang dengan tingkat akurasi yang memadai. Di antara berbagai algoritma machine learning, metode ensemble berbasis regresi seperti Random Forest Regression telah menunjukkan performa yang baik dalam menangani data utilisasi resource berbasis time-series karena ketahanannya terhadap noise dan kemampuannya menangkap hubungan non-linear antar fitur tanpa memerlukan normalisasi data secara ekstensif[4].

Penelitian ini bertujuan untuk mengkaji penerapan algoritma Random Forest Regression untuk memprediksi utilisasi sumber daya pada lingkungan private cloud berdasarkan data beban kerja historis dari dataset GWA-T-12 Bitbrains[3]. Target prediksi mencakup empat metrik sumber daya utama: CPU usage, memory usage, disk write throughput, dan network transmitted throughput.

METODOLOGI

Desain Penelitian

Penelitian ini menggunakan desain penelitian kuantitatif yang bersifat eksperimental berdasarkan analisis data historis. Penelitian ini berfokus pada pembangunan dan evaluasi model prediktif yang mampu memperkirakan utilisasi sumber daya berdasarkan pola historis[5]. Alur penelitian terdiri dari empat tahapan utama: pengumpulan data, praproses data, pelatihan model, dan evaluasi model.

Sumber Data

Penelitian ini menggunakan dataset yang terdiri dari 1.250 VM terhubung ke storage area network (SAN) berkecepatan tinggi[3]. Data dikumpulkan dengan interval pencatatan 5 menit, menghasilkan sekitar 8.640 observasi per VM. Tiga VM sampel digunakan sebagai representasi beban kerja yang beragam (VM 99, VM 987, VM 988). Setiap file CSV berisi 11 kolom: Timestamp, CPU cores, CPU capacity provisioned (MHz), CPU usage (MHz), CPU usage (%), Memory provisioned (KB), Memory usage (KB), Disk read throughput (KB/s), Disk write throughput (KB/s), Network received throughput (KB/s), dan Network transmitted throughput (KB/s).

Tabel 1. Deskripsi Dataset

Atribut	VM 99	VM 987	VM 988
Jumlah observasi	8.619	8.640	8.640
CPU cores	1	2	2
CPU capacity (MHz)	2.600	4.800	4.800
Memory provisioned	2 GB	4 GB	4 GB
Rata-rata CPU usage (%)	0,28	4,68	1,16
Rata-rata Memory usage	~120 MB	~474 MB	~214 MB

Praproses Data

Sebelum pelatihan model, dataset melalui tahap praproses yang mencakup konversi timestamp ke format datetime serta konstruksi fitur berbasis riwayat waktu. Fitur yang dibangun meliputi: (1) fitur lag (lag_1 hingga lag_12), yaitu nilai target pada 12 periode sebelumnya setara 1 jam riwayat; (2) rolling statistics berupa rolling mean, rolling standard deviation, dan rolling maximum dengan window 3, 6, dan 12 periode (15, 30, dan 60 menit); (3) fitur delta (diff_1), yaitu selisih antara lag_1 dan lag_2 untuk menangkap arah tren; serta (4) fitur waktu berupa jam dalam sehari, hari dalam seminggu, dan indikator akhir pekan. Khusus untuk Disk Write Throughput, fitur Disk Read Throughput disertakan sebagai prediktor tambahan. Pembagian data menggunakan rasio 80:20 berbasis urutan waktu (bukan pengacakan).

Algoritma Machine Learning

Penelitian ini menggunakan Random Forest Regression sebagai algoritma utama[4]. Untuk target dengan distribusi sangat condong (skewed) seperti Disk Write Throughput dan Network Transmitted Throughput, transformasi log (log1p) diterapkan pada variabel

target sebelum pelatihan, dan hasil prediksi dikembalikan ke skala asli menggunakan transformasi invers (expm1) sebelum evaluasi. Konfigurasi hyperparameter disajikan pada Tabel 2.

Tabel 2. Konfigurasi Hyperparameter Model Random Forest Regression

Target Resource	n_estimators / max_depth	Transformasi Target
CPU Usage (%)	150 / 20	Tidak ada
Memory Usage (KB)	150 / 20	Tidak ada
Disk Write Throughput (KB/s)	150 / 15	log1p / expm1
Network Transmitted (KB/s)	150 / 10	log1p / expm1

Evaluasi Model

Performa prediksi model dievaluasi menggunakan tiga metrik evaluasi regresi yang umum digunakan dalam penelitian machine learning, yaitu Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan koefisien determinasi (R^2).

MAE mengukur rata-rata besaran selisih absolut antara nilai prediksi dan nilai aktual, tanpa mempertimbangkan arah kesalahan. Metrik ini memberikan gambaran langsung tentang seberapa besar rata-rata deviasi prediksi dalam satuan yang sama dengan variabel target, sehingga mudah diinterpretasikan secara intuitif.

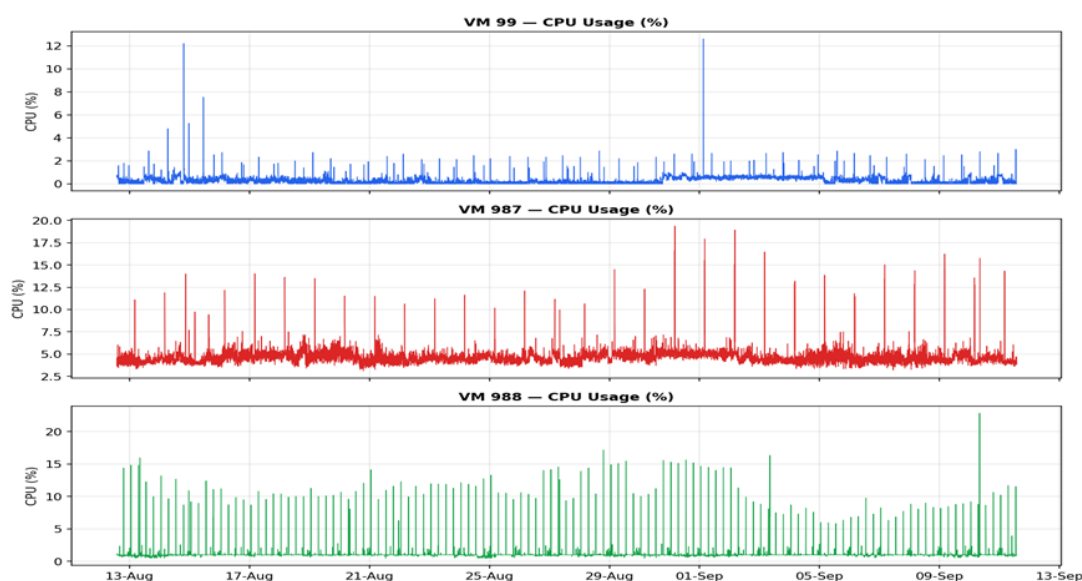
RMSE dihitung sebagai akar kuadrat dari rata-rata kuadrat selisih antara nilai prediksi dan nilai aktual. Dibandingkan dengan MAE, RMSE memberikan penalti yang lebih besar terhadap kesalahan prediksi yang besar (outlier error), sehingga lebih sensitif terhadap prediksi yang meleset jauh dari nilai aktual. Nilai RMSE yang rendah mengindikasikan bahwa model tidak hanya akurat secara rata-rata, tetapi juga konsisten tanpa kesalahan ekstrem yang besar.

R^2 (koefisien determinasi) mengukur proporsi variansi pada data aktual yang berhasil dijelaskan oleh model. Nilai R^2 berkisar antara 0 hingga 1 pada kondisi ideal, di mana nilai mendekati 1 menunjukkan bahwa model mampu menjelaskan sebagian besar variansi data, sementara nilai mendekati 0 atau negatif menunjukkan bahwa model tidak lebih baik dari sekadar menebak nilai rata-rata. Dalam penelitian ini, R^2 digunakan sebagai metrik utama untuk membandingkan performa model antar jenis resource dan antar VM, karena sifatnya yang tidak bergantung pada satuan pengukuran sehingga perbandingan lintas resource (CPU, memori, disk, jaringan) dapat dilakukan secara langsung.

HASIL DAN PEMBAHASAN

Statistik Deskriptif Utilisasi Sumber Daya

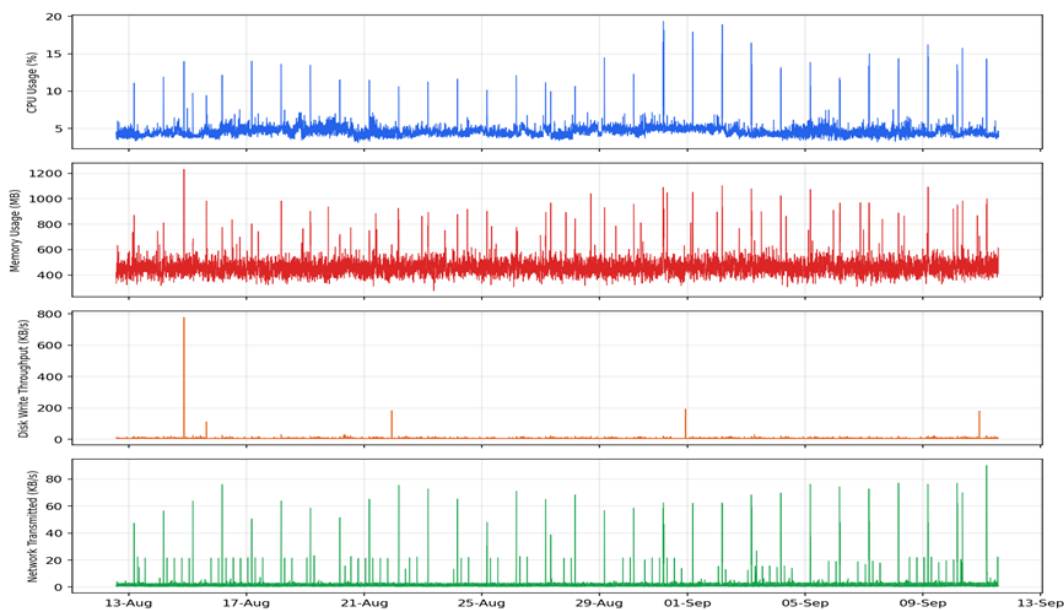
Hasil eksplorasi data menunjukkan bahwa ketiga VM memiliki profil beban kerja yang sangat berbeda. Gambar 1 menyajikan pola utilisasi CPU dari ketiga VM selama periode pengamatan satu bulan.



Gambar 1. Pola Utilisasi CPU per VM

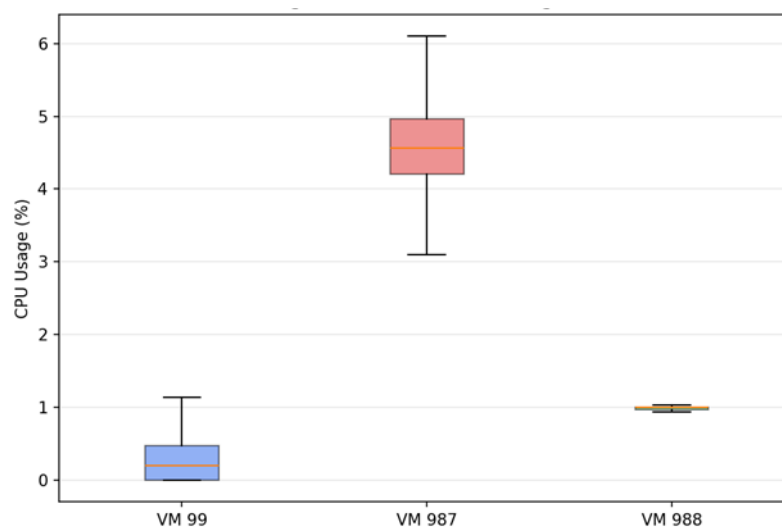
Dari Gambar 1 terlihat bahwa VM 99 beroperasi dalam kondisi beban kerja sangat ringan dengan spike yang jarang terjadi, VM 987 memiliki baseline CPU relatif tinggi dan stabil dengan spike periodik yang konsisten, sedangkan VM 988 menunjukkan pola spike

yang paling sering dengan variasi amplitudo besar. Gambar 2 menyajikan keempat dimensi utilisasi sumber daya untuk VM 987 sebagai representasi VM dengan beban kerja paling kompleks.



Gambar 2. Empat Dimensi Utilisasi Sumber Daya VM 987

Gambar 2 mengungkapkan bahwa lonjakan pada keempat dimensi resource cenderung terjadi secara bersamaan, mengindikasikan adanya proses terjadwal yang secara simultan memengaruhi CPU, memori, disk, dan jaringan. Temuan ini mendukung keputusan untuk menyertakan fitur Disk Read Throughput sebagai prediktor tambahan dalam model prediksi Disk Write Throughput. Gambar 3 memperlihatkan perbandingan distribusi CPU Usage antar ketiga VM.



Gambar 3. Perbandingan Distribusi CPU Usage Antar VM

Performa Prediksi Model

Hasil evaluasi model Random Forest Regression untuk seluruh kombinasi VM dan target resource disajikan pada Tabel 3. Evaluasi dilakukan pada data uji yang mencakup 20% data terakhir secara kronologis (sekitar 1.722–1.726 observasi per VM), sehingga model diuji pada kondisi yang belum pernah dilihat selama pelatihan dan mencerminkan kemampuan generalisasi model terhadap data masa depan.

Secara umum, hasil pada Tabel 3 menunjukkan variasi performa yang cukup signifikan antar jenis resource maupun antar VM. Beberapa kombinasi mencatatkan nilai R^2 yang tinggi, mengindikasikan kemampuan prediksi yang kuat, sementara kombinasi lainnya menunjukkan nilai R^2 yang lebih rendah, yang mencerminkan tantangan inherent dalam memprediksi resource dengan pola beban kerja yang lebih tidak teratur. Perbedaan nilai MAE dan RMSE antar resource juga mencerminkan perbedaan skala satuan masing-masing

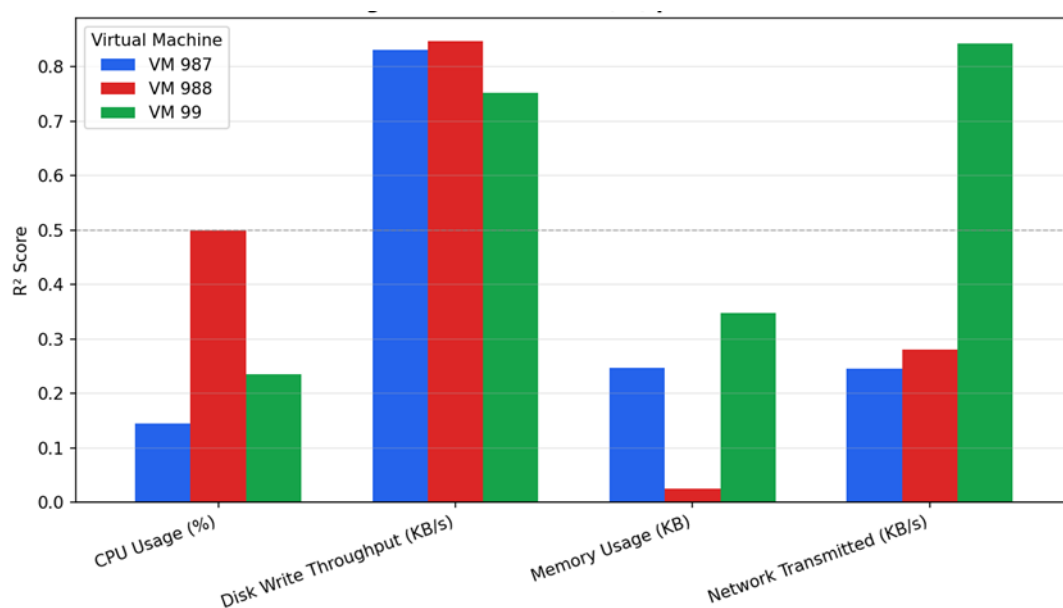
metrik (persen untuk CPU, kilobyte untuk memori, dan kilobyte per detik untuk disk dan jaringan), sehingga perbandingan performa lintas resource sebaiknya mengacu pada nilai R^2 yang bersifat bebas satuan.

Tabel 3. Hasil Evaluasi Model Random Forest Regression per VM dan Target Resource

VM	Target Resource	MAE	RMSE	R^2
VM 99	CPU Usage (%)	0,1479	0,2757	0,2350
VM 99	Memory Usage (KB)	22.854,87	46.889,58	0,3476
VM 99	Disk Write Throughput (KB/s)	0,6747	2,8322	0,7522
VM 99	Network Transmitted (KB/s)	0,0228	0,0564	0,8431
VM 987	CPU Usage (%)	0,4136	0,9053	0,1444
VM 987	Memory Usage (KB)	48.150,13	69.085,61	0,2466
VM 987	Disk Write Throughput (KB/s)	1,2994	2,0050	0,8310
VM 987	Network Transmitted (KB/s)	0,9328	5,5052	0,2452
VM 988	CPU Usage (%)	0,2099	0,8479	0,4996
VM 988	Memory Usage (KB)	39.080,58	67.050,07	0,0253
VM 988	Disk Write Throughput (KB/s)	0,9804	5,9920	0,8463
VM 988	Network Transmitted (KB/s)	0,6168	4,2979	0,2810

Gambar 4 menyajikan perbandingan nilai R^2 secara visual untuk seluruh kombinasi VM dan target resource dalam bentuk diagram batang berkelompok. Sumbu vertikal menunjukkan nilai R^2 , sementara sumbu horizontal mengelompokkan hasil berdasarkan jenis resource. Garis putus-putus horizontal pada nilai $R^2 = 0,5$ digunakan sebagai referensi ambang batas performa sedang, sehingga memudahkan identifikasi kombinasi mana yang berada di atas atau di bawah ambang tersebut secara sekilas.

Dari Gambar 4 terlihat dengan jelas bahwa Disk Write Throughput secara konsisten menunjukkan nilai R^2 tertinggi pada ketiga VM, dengan seluruh batang berada di atas nilai 0,75. Hal ini mengindikasikan bahwa model berhasil memprediksi throughput tulis disk dengan tingkat konsistensi yang baik lintas VM yang berbeda. Sebaliknya, CPU Usage dan Memory Usage menampilkan nilai R^2 yang lebih rendah dan bervariasi antar VM, mencerminkan tingkat kesulitan yang lebih tinggi dalam memprediksi kedua resource tersebut hanya dari riwayat penggunaan historis. Network Transmitted Throughput menunjukkan hasil yang paling bervariasi, dengan VM 99 mencatatkan nilai R^2 yang sangat tinggi sementara VM 987 dan VM 988 berada pada kisaran yang lebih rendah, mengindikasikan bahwa pola jaringan pada ketiga VM ini memiliki karakteristik yang berbeda secara mendasar.



Gambar 4. Perbandingan Nilai Model R^2 per Resource dan VM

Dari Tabel 3 dan Gambar 4, Disk Write Throughput dan Network Transmitted menunjukkan nilai R^2 yang relatif lebih tinggi dibandingkan CPU Usage dan Memory Usage. Pola ini konsisten ditemukan pada ketiga VM yang diuji, mengindikasikan bahwa kedua resource tersebut memiliki karakteristik yang lebih dapat diprediksi menggunakan pendekatan berbasis riwayat historis dibandingkan

CPU dan memori. Hal ini dapat dikaitkan dengan sifat beban kerja I/O disk dan jaringan yang cenderung mengikuti pola terjadwal atau dipicu oleh event yang berulang secara periodik, sehingga informasi dari periode sebelumnya (fitur lag dan rolling statistics) lebih efektif menangkap polanya. Sebaliknya, beban kerja CPU dan memori lebih bersifat reaktif terhadap permintaan pengguna atau proses internal aplikasi yang dinamikanya lebih sulit ditangkap hanya dari riwayat penggunaan itu sendiri.

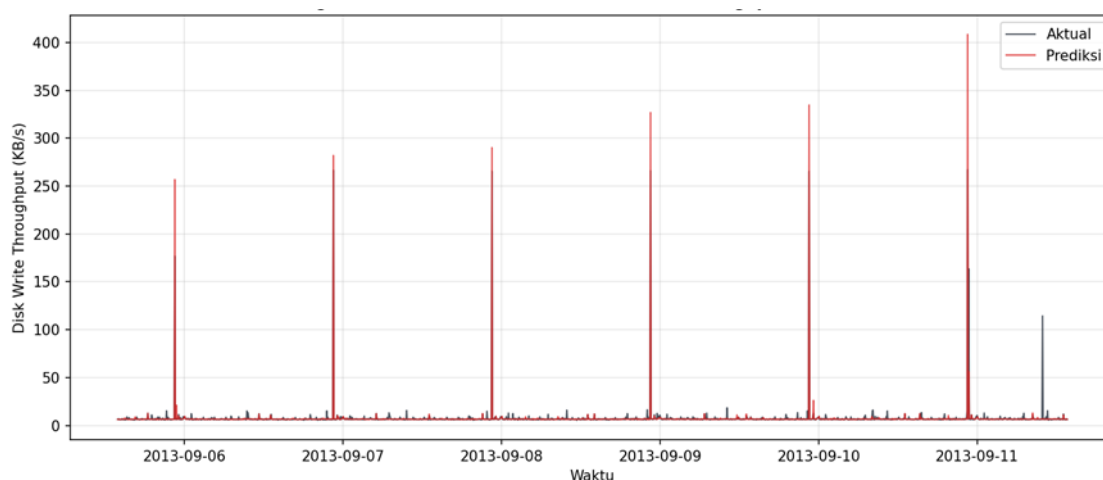
Rata-rata nilai R^2 per jenis resource disajikan pada Tabel 4 untuk memberikan gambaran agregat performa model secara lintas VM. Nilai rata-rata ini dihitung dari hasil ketiga VM (VM 99, VM 987, VM 988) dan dilengkapi dengan nilai R^2 minimum dan maksimum untuk menggambarkan rentang variasi performa antar VM pada masing-masing jenis resource.

Tabel 4. Rata-rata Nilai R^2 per Jenis Resource (dari 3 VM)

Target Resource	R^2 Rata-rata	R^2 Minimum	R^2 Maksimum
CPU Usage (%)	0,2930	0,1444	0,4996
Disk Write Throughput (KB/s)	0,8098	0,7522	0,8463
Memory Usage (KB)	0,2065	0,0253	0,3476
Network Transmitted (KB/s)	0,4564	0,2452	0,8431

Gambar 5 menyajikan perbandingan nilai aktual dan prediksi untuk Disk Write Throughput pada VM 988 ($R^2 = 0,846$) selama periode data uji. Grafik ini menampilkan dua kurva yang diplot secara bersamaan: kurva hitam merepresentasikan nilai aktual throughput tulis disk yang tercatat dari dataset, sementara kurva merah merepresentasikan hasil prediksi model Random Forest Regression.

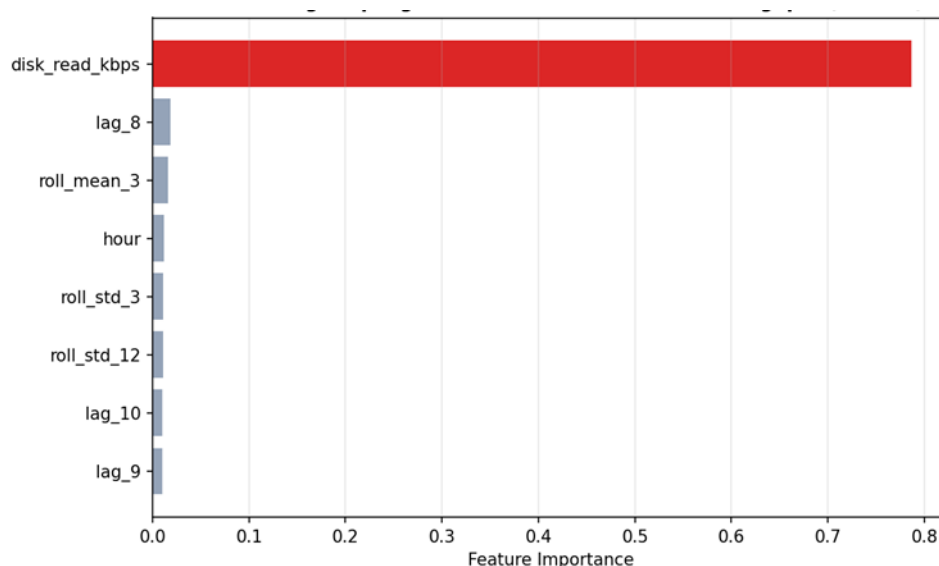
Dari Gambar 5 terlihat bahwa model berhasil menangkap pola lonjakan (spike) Disk Write Throughput yang terjadi secara periodik, di mana kemunculan spike pada kurva prediksi secara umum bertepatan dengan spike pada data aktual. Hal ini menunjukkan bahwa model mampu mengidentifikasi pola temporal berulang yang mengindikasikan kapan aktivitas tulis disk intensif akan terjadi. Namun demikian, terdapat perbedaan pada magnitude spike di beberapa titik, di mana prediksi model cenderung meremehkan (underestimate) nilai puncak yang sangat tinggi. Kondisi ini merupakan karakteristik umum dari model ensemble berbasis averaging seperti Random Forest, yang secara alami menghasilkan prediksi yang lebih konservatif karena merata-ratakan kontribusi banyak decision tree. Meskipun demikian, untuk keperluan alokasi resource secara proaktif, kemampuan model dalam mengantisipasi waktu terjadinya lonjakan resource sudah memberikan nilai praktis yang signifikan, meskipun estimasi besarannya belum sempurna.



Gambar 5. Perbandingan Nilai Aktual vs Prediksi - Disk Write Throughput VM 988 ($R^2 = 0,846$)

Analisis Feature Importance

Gambar 6 menyajikan delapan fitur paling berpengaruh dalam model prediksi Disk Write Throughput untuk VM 988 berdasarkan ukuran feature importance dari Random Forest.



Gambar 6. Delapan Fitur Paling Berpengaruh - Prediksi Disk Write Throughput VM 988

Gambar 6 menunjukkan bahwa fitur `disk_read_kbps` mendominasi kontribusi prediksi dengan nilai importance sebesar 0,786, jauh melampaui fitur-fitur lainnya. Pola serupa juga ditemukan pada VM 99 (`disk_read_kbps` = 0,475) dan VM 987 (`disk_read_kbps` = 0,220), mengonfirmasi bahwa korelasi antara aktivitas baca dan tulis disk merupakan informasi prediktif paling signifikan untuk resource ini.

Interpretasi Performa Model

Disk Write Throughput memperoleh nilai R^2 tertinggi dengan rata-rata 0,81 lintas tiga VM. Peningkatan performa ini terutama disebabkan oleh penyertaan fitur Disk Read Throughput, yang terbukti menjadi prediktor terkuat karena kedua metrik mencerminkan aktivitas I/O yang dipicu oleh proses yang sama. Network Transmitted Throughput memperoleh rata-rata R^2 sebesar 0,46 dengan variasi besar antar VM (0,25-0,84); pada VM 99 performa sangat baik ($R^2 = 0,84$) dengan fitur `lag_12` sebagai prediktor terkuat, mengindikasikan pola periodik kuat dengan periode sekitar 1 jam. CPU Usage memperoleh rata-rata R^2 sebesar 0,29, yang tergolong sedang-lemah namun dimaklumi mengingat karakteristik spike yang bersifat bursty pada data workload nyata. Memory Usage memperoleh rata-rata R^2 terendah (0,21), dengan VM 988 mencatatkan nilai yang sangat rendah ($R^2 = 0,025$), mengindikasikan bahwa pola penggunaan memori lebih dipengaruhi dinamika internal aplikasi yang tidak dapat diprediksi hanya dari riwayat penggunaan itu sendiri.

Perbandingan dengan Pendekatan Alokasi Konvensional

Dibandingkan dengan metode alokasi sumber daya konvensional berbasis ambang batas statis, pendekatan prediksi berbasis machine learning menawarkan keunggulan fundamental: kemampuan mengantisipasi perubahan kebutuhan sumber daya sebelum terjadi. Pada konteks Disk Write Throughput misalnya, model yang mampu memprediksi lonjakan I/O disk beberapa periode ke depan memungkinkan sistem melakukan pre-provisioning sebelum lonjakan mencapai puncaknya, menghindari bottleneck yang berdampak pada performa aplikasi.

Temuan Kunci: Korelasi Antar-Resource

Salah satu temuan paling signifikan adalah peran kritis korelasi antar-resource dalam meningkatkan akurasi prediksi. Penambahan fitur Disk Read Throughput menghasilkan peningkatan R^2 yang drastis dari mendekati 0 menjadi lebih dari 0,75 pada semua VM. Temuan ini mengimplikasikan bahwa sistem prediksi resource untuk private cloud perlu mempertimbangkan hubungan antar-resource yang secara fungsional saling berkaitan (pendekatan multivariate), bukan hanya memodelkan setiap resource secara terpisah (univariate).

Implikasi Praktis

Model prediksi berbasis Random Forest Regression dapat diintegrasikan ke dalam sistem manajemen private cloud sebagai komponen decision-support untuk perencanaan resource yang lebih proaktif. Model Disk Write dan Network Transmitted yang menunjukkan performa prediksi cukup kuat dapat dipertimbangkan untuk skenario auto-scaling berbasis prediksi, sementara model CPU dan Memory memerlukan pengembangan lebih lanjut.

KESIMPULAN DAN SARAN

Penelitian ini telah menyajikan pendekatan optimalisasi alokasi sumber daya pada arsitektur private cloud melalui penerapan algoritma Random Forest Regression untuk memprediksi utilisasi CPU, memori, disk write throughput, dan network transmitted throughput. Hasil eksperimen menunjukkan bahwa model mampu memprediksi Disk Write Throughput dengan rata-rata R^2 sebesar 0,81 dan Network Transmitted Throughput dengan rata-rata R^2 sebesar 0,46 lintas tiga VM yang diuji. Peningkatan signifikan pada prediksi Disk Write Throughput diperoleh melalui penyertaan fitur Disk Read Throughput sebagai prediktor tambahan, mengungkap pentingnya pemodelan korelasi antar-resource dalam sistem prediksi workload cloud. Sementara itu, prediksi CPU Usage dan Memory Usage masih tergolong sedang-lemah, mengindikasikan perlunya pengembangan fitur atau algoritma yang lebih canggih. Penelitian selanjutnya disarankan untuk: (1) mengeksplorasi implementasi real-time pada testbed private cloud; (2) membandingkan performa Random Forest dengan Gradient Boosting dan LSTM; (3) memperluas evaluasi pada lebih banyak VM dengan beban kerja lebih beragam; serta (4) mengintegrasikan fitur kontekstual tambahan untuk meningkatkan akurasi prediksi CPU dan Memory.

DAFTAR PUSTAKA

Jurnal

- [1] S. Shen, V. van Beek, and A. Iosup, "Statistical Characterization of Business-Critical Workloads Hosted in Cloud Datacenters," in Proc. 15th IEEE/ACM Int. Symp. Cluster, Cloud Grid Comput. (CCGrid), 2015, pp. 465-474.
- [2] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5-32, 2001.
- [3] R. Moreno-Vozmediano et al., "Orchestrating the Deployment of High Performance Computing Workloads on Hybrid Clouds," *Future Gener. Comput. Syst.*, vol. 112, pp. 528-541, 2020.
- [4] Zulherry, A. (2026). Implementation of Multi-Room Computer Laboratory Network Infrastructure Based on Star Topology in an Educational Environment. *Hanif Journal of Information Systems*, 3(2), 76-81.
- [5] Vianingrum, M. H., & Zulherry, A. (2026). Implementasi Sistem Pendukung Keputusan Untuk Menganalisis Tingkat Kepuasan Pengguna E-Learning Menggunakan Metode End User Computing Satisfaction (EUCS) Di SMK Multi Karya. *Hello World Jurnal Ilmu Komputer*, 5(1), 77-84.
- [6] Lubis, I. Q., & Zulherry, A. (2026). Implementasi Keamanan Website Dari Serangan Cross Site Request Forgery Menggunakan Algoritma HMAC-SHA256 Pada Framework Laravel. *Hello World Jurnal Ilmu Komputer*, 5(1), 47-58.
- [7] Zulherry, A., Sari, I. P., & Basri, M. (2026). Perancangan dan Implementasi Segmentasi LAN pada Infrastruktur Jaringan Skala Menengah. *Hello World Jurnal Ilmu Komputer*, 4(4).
- [8] Zulherry, A., Riadi, I., & Umar, R. (2026). Anomaly Detection in Cloud Device-Based Information Technology Infrastructure Using Isolation Forest Algorithm. *Journal Of Informatics And Telecommunication Engineering* 9 (2)
- [9] Zulherry, A., Gunawan, M., & Basri, M. (2025). Design and Implementation of Multi-Segment LAN Infrastructure for Computer Laboratories. *Tsabit Journal of Computer Science*, 2(2), 110-118.
- [10] Zulherry, A., Basri, M., & Yusnandar, W. (2025). Implementation of Private Cloud Infrastructure as a Service Using Proxmox VE Based on the Network Development Life Cycle (NDLC) Approach. *Altafani: Jurnal Pengabdian Masyarakat*, 1(1).
- [11] Zulherry, A., & Gunawan, M. (2025). Development of an Android-Based Smart Health Monitoring Device for Heartbeat Detection. *Al'adzkiya International of Computer Science and Information Technology (AIOCSIT) Journal*, 6(2), 75-79.
- [12] Basri, M., & Zulherry, A. (2025). Analysis of the Impact of Gambling and Online Loans in the Perspective of Informatics, Islam, and Kemuhammadiyah. *Ar-Rasyid: Jurnal Pendidikan Agama Islam*, 5(1), 65-73.
- [13] Asadel, A., & Zulherry, A. (2025). Detecting Zero-Width Characters Obfuscated in Phishing URLs using the XGBOOST Algorithm. *Hanif Journal of Information Systems* 3 (1), 43-53
- [14] Zulherry, A., Sari, I.P., & Basri, M. (2025). Perancangan Aplikasi Monitoring Kehadiran Pegawai Menggunakan RFID. *sudo Jurnal Teknik Informatika* 4 (4), 378-384
- [15] Ichsan, A., Zulherry, A., Lubis, T. A., & Shahnaz, B. A. Z. (2025). Utilization of Mobile Applications to Speed Up The Search for Android-Based Index Places. *IJATCoS: Indonesian Journal of Applied Technology. Computer and Science*, 2(1).
- [16] Amada, E.P., & Zulherry, A. (2025). Klasterisasi Minat Dan Bakat Siswa Menggunakan Metode X-Means Berbasis Web: Studi Kasus SMA Negeri 1 Hamparan Perak. *sudo Jurnal Teknik Informatika* 4 (4), 276-283.

- [17]Zulherry, A., Ramadhani, F., & Satria, A. (2024). Klasifikasi Data Tracer Study Dengan Pemanfaatan Data Mining Menggunakan Algoritma Support Vector Machine dan Neural Network. *Portal Riset dan Inovasi Sistem Perangkat Lunak* 2 (1), 45-54
- [18]Zulherry, A. (2023). Decision making for network security with simple additive weighting method. *Journal of Intelligent Decision Support System (IDSS)*, 6(3), 155-159.
- [19]Zulherry, A., Gunawan, T. S., & Wanayumini, W. (2021). Analisis Hasil Pendukung Keputusan Mendapatkan Rumah Dinas Perusahaan Menggunakan Metode Analytical Hierarchy Process (AHP) dan Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). *Jurnal Media Informatika Budidarma*, 5(2), 695-704.
- [20]Sari, I.P., Zulherry, A., Basri, M., & Hayani W. (2025). Pembelajaran Pemrograman berbasis Machine Learning sebagai Upaya Peningkatan Computational Thinking. *Jurnal Penelitian, Pendidikan dan Pengajaran: JPPP* 6 (3), 245-250
- [21]Sari, I.P., & Batubara, I.H. (2021). User Interface Information System for Using Account Services (Joint Account) WEB-Based. *International Journal of Economic, Technology and Social Sciences (Injects)* 2 (2), 462-469
- [22]Ramadhani, F., & Sari, I.P. (2021). Pemanfaatan Aplikasi Online dalam Digitalisasi Pasar Tradisional di Medan. *Prosiding Seminar Nasional Kewirausahaan* 2 (1), 806-811
- [23]Sari, I.P., & Alfari, F. (2024). Perancangan Sistem Aplikasi Pendataan Membership Gym Menggunakan Metode Unified Software Development Process (USDP) Berbasis Web. *Hello World Jurnal Ilmu Komputer* 3 (1), 37-48
- [24]Sari, I.P., Jannah, A., Meuraxa, A.M., Syahfitri, A., & Omar, R. (2022). Perancangan Sistem Informasi Penginputan Database Mahasiswa Berbasis Web. *Hello World Jurnal Ilmu Komputer* 1 (2), 106-110
- [25]Satria, A., Ramadhani, F., & Sari, I.P. (2023). Rancang Bangun Sistem Informasi Penerimaan Peserta Didik Baru (PPDB) Sekolah Menengah Kejuruan Telkom 2 Medan Menggunakan Codeigniter. *Wahana Jurnal Pengabdian kepada Masyarakat* 2 (1), 23-31
- [26]Sari, I.P., Azzahrah, A., Qathrunada, I.F., Lubis, N., & Anggraini, T. (2022). Perancangan sistem absensi pegawai kantor secara online pada website berbasis HTML dan CSS. *Blend sains jurnal teknik* 1 (1), 8-15
- [27]Hariani, P.P., Sari, I.P., & Batubara, I.H. (2021). Android-Based Financial Statement Presentation Model. *JURNAL TARBIYAH* 28 (2), 1-16
- [28]Sari, I.P., Syahputra, A., Zaky, N., Sibuea, R.U., & Zakhir, Z. (2022). Perancangan sistem aplikasi penjualan dan layanan jasa laundry sepatu berbasis website. *Blend sains jurnal teknik* 1 (1), 31-37
- [29]Sari, I.P., Batubara, I.H., Ramadhani, F., & Wardani, S. (2022). Perancangan Sistem Antrian pada Wahana Hiburan dengan Metode First In First Out (FIFO). *Sudo Jurnal Teknik Informatika* 1 (3), 116-123
- [30]Ramadhani, F., Satria, A., & Sari, I.P. (2022). Aplikasi internet berbasis website sebagai E-Commerce penjualan komponen sport car. *Blend Sains Jurnal Teknik* 1 (2), 69-75
- [31]Sari, I.P., Ramadhani, F., Satria, A., Apdilah, D., & Basri, M. (2023). Rancangan UI/UX Aplikasi Analytics pada Toko Online Wao Sneakers Menggunakan Figma Berbasis Mobile. *Factory Jurnal Industri, Manajemen dan Rekayasa Sistem Industri* 1 (3), 93-101
- [32]Sari, I.P., Al-Khowarizmi, A., & Batubara, I.H. (2021). Cluster Analysis Using K-Means Algorithm and Fuzzy C-Means Clustering For Grouping Students' Abilities In Online Learning Process. *Journal of Computer Science, Information Technology and Telecommunication Engineering* 2 (1), 139-144
- [33]Hutasuhut, B.K., Sari, I.P., & Al-Khowarizmi, A. (2023). Analysis the Effect of Digitalization and Technology on Web-Based Entrepreneurship. *Journal of Computer Science, Information Technology and Telecommunication Engineering* 4 (1), 350-354
- [34]Sitompul, D.N., Rahmatika, A., & Sari, I.P. (2023). Application of The Sales and Purchase Program Using The Rapid Application Development Model. *Al'adzkiya International of Computer Science and Information Technology (AIOCSIT) Journal* 4 (1), 6-16
- [35]Sari, I.P., & Batubara, I.H. (2021). Perancangan Sistem Informasi Laporan Keuangan Pada Apotek Menggunakan Algoritma K-NN. *Seminar Nasional Teknologi Edukasi dan Humaniora (SiNTeSa)* (1).
- [36]Ramadhani, F., Satria, A., & Sari, I.P. (2023). Implementasi Metode Fuzzy K-Nearest Neighbor dalam Klasifikasi Penyakit Demam Berdarah. *Hello World Jurnal Ilmu Komputer* 2 (2), 58-62
- [37]Sari, I.P., Batubara, I.H., & Basri, M. (2022). Implementasi Internet of Things Berbasis Website dalam Pemesanan Jasa Rumah Service Teknisi Komputer dan Jaringan Komputer. *Blend Sains Jurnal Teknik* 1 (2), 157-163
- [38]Sari, I.P., & Ramadhani, F. (2021). Pengaruh Teknologi Informasi Terhadap Kewirausahaan Pada Aplikasi Perancangan Jual Beli Jamu Berbasis WEB. *Prosiding Seminar Nasional Kewirausahaan* 2 (1), 874-878
- [39]Rangkuti, A.R.S., & Sari, I.P. (2026). Penggunaan Fitur Warna Cielab Dalam Segmentasi Kematangan Buah Kakao Dengan Algoritma K-Means Clustering. *Hello World Jurnal Ilmu Komputer* 5 (1), 36-46

- [40] Sari, I.P., Basri, M., & Syafrayani, P.R. (2025). Implementasi Sistem Aplikasi Pengolahan Teks pada Gambar Menggunakan Modifikasi Metode LSB dan ROT13. *Hello World Jurnal Ilmu Komputer* 4 (2), 90-97
- [41] Sari, I.P., Fadlil, A., & Sutikno, T. (2025). Islamic sound recognition using MFCC and SVM: Case study on takbir and sholawat. *Journal of Computer Networks, Architecture and High Performance Computing* 7 (3), 825-839
- [42] Sari, I.P., Azis, Z., Hasibuan, A.R. (2025). Analysis Of Ron 92 Oil Based On Morphology And Histogram Techniques. *BAREKENG: Jurnal Ilmu Matematika dan Terapan* 19 (2), 1341-1352
- [43] Aqta, D.P., & Sari, I.P. (2025). Evaluasi Keefektifan Teknik Morfologi dan Histogram Citra Digital pada Minyak RON 92 di SPBU Pertamina Medan Tembung. *Hello World Jurnal Ilmu Komputer* 3 (4), 161-170
- [44] Sari, I.P., Sulaiman, O.K., & Apdilah, D. (2025). Penerapan Sistem Pakar Diagnostik Penyakit Kelapa Sawit sebagai Solusi bagi Petani dalam Meningkatkan Produktivitas Perkebunan. *CivicAction: Jurnal Pengabdian dan Inovasi Masyarakat* 1 (1), 8-17

Buku

- [45] Indah Purnama Sari. *Algoritma dan Pemrograman*. Medan: UMSU Press, 2023, pp. 290.
- [46] Janner Simarmata Arsan Kumala Jaya, Syarifah Fitrah Ramadhani, Niel Ananto, Abdul Karim, Betrisandi, Muhammad Ilham Alhari, Cucut Susanto, Suardinata, Indah Purnama Sari, Edson Yahuda Putra. *Komputer dan Masyarakat*. Medan: Yayasan Kita Menulis, 2024, pp.162.
- [47] Mahdianta Pandia, Indah Purnama Sari, Alexander Wirapraja Fergie Joanda Kaunang, Syarifah Fitrah Ramadhani Stenly Richard Pungus, Sudirman, Suardinata Jimmy Herawan Moedjahedy, Elly Warni, Debby Erce Sondakh. *Pengantar Bahasa Pemrograman Python*. Medan : Yayasan Kita Menulis, 2024, pp.180
- [48] Zelvi Gustiana Arif Dwinanto, Indah Purnama Sari, Janner Simarmata Mahdianta Pandia, Supriadi Syam, Semmy Wellem Taju Fitrah Eka Susilawati, Asmah Akhriana, Rolly Junius Lontaan Fergie Joanda Kaunang. *Perkembangan Teknologi Informatika*. Medan: Yayasan Kita Menulis, 2024, pp.158
- [49] Indah Purnama Sari. *Buku Ajar Pemrograman Internet Dasar*. Medan: UMSU Press, 2022, pp. 300.
- [50] Indah Purnama Sari. *Buku Ajar Rekayasa Perangkat Lunak*. Medan: UMSU Press, 2021, pp. 228.
- [51] Purba, A. M. V., Pradipta, R., Zulherry, A., Sari, I. P., Ansya, Y. A. U., Basri, M., ... & Darnilasari, A. (2026). *Pengetahuan Dasar Komputer*. Yayasan Kita Menulis.
- [52] Kaban, R., Rachman, A. I., Sari, I. P., Zulherry, A., Basri, M., Fadilah, N., ... & Ahyuna, A. (2026). *Pengantar Artificial Intelligence, Machine Learning, dan Big Data*. Yayasan Kita Menulis.
- [53] Tarigan, I. M. B., Sari, I. P., Zulherry, A., Ansya, Y. A. U., Laman, E. G., Saputra, E. M., ... & Damayanti, E. (2026). *Fondasi Matematika untuk Ilmu Komputer dan Implementasinya dalam Pendidikan*. Yayasan Kita Menulis.
- [54] Irfan, A. P., Susanto, C., Sari, I. P., Zulherry, A., Ansya, Y. A. U., Dewi, I. K., ... & Mubarak, R. (2026). *Sistem Informasi Modern untuk Era Digital*. Yayasan Kita Menulis.
- [55] Wahyuddin S, Indah Purnama Sari, Samuel Hasudungan Tampubolon, Andi Zulherry, Butsiarah, Taufik, George Morris William Tangka Rahmad Bahri, Mhd. Basri, Raemon Syaljumairi, Robbi Rahim Muharman Lubis, Baso Ali, Sirmayanti, Janner Simarmata. *Cloud Computing: Referensi Lengkap*. Medan: Yayasan Kita Menulis, 2025, pp.248.