

Analisis Trade-off Komunikasi-Komputasi pada Federated Averaging untuk Klasifikasi Penyakit Jantung

Eldwin Fikhar Ananda
Fakultas Informatika
Universitas Telkom
Bandung

eldwinfikhar@students.telkomuniv
ersity.ac.id

Aji Gautama Putrada
Fakultas Informatika
Universitas Telkom
Bandung

ajigps@telkomuniversity.ac.id

Ayu Qatrunnada Istiqfari
Fakultas Informatika
Universitas Telkom
Bandung

ayuatrunnada@telkomuniversity.a
c.id

Klasifikasi penyakit jantung menggunakan Federated Learning (FL) mampu menjadi solusi untuk meningkatkan privasi di dunia medis. Sayangnya, penerapan FL terkendala oleh tingginya *communication cost* yang harus dioptimalkan dengan turut mempertimbangkan *computational load* yang rendah. Penelitian ini mengajukan sebuah arsitektur FL yang memadukan Federated Averaging (FedAvg) dengan model Multi-Layer Perceptron (MLP) untuk melakukan klasifikasi penyakit jantung menggunakan dataset UCI Heart Disease. Arsitektur ini telah dilibatkan dalam pelatihan FL yang dilakukan hingga akurasi mendekati skenario pelatihan terpusat (*centralized*) sebagai titik *threshold* dengan *local epoch* (E) yang bervariasi. Terdapat dua metrik yang diperhitungkan dalam penelitian ini, yaitu *communication cost* dan *computational load* yang akan dimodelkan menggunakan regresi linear. Selanjutnya, telah ditentukan juga titik optimal antara keduanya melalui metode Weighted Sum Model (WSM). Hasil pengujian menunjukkan bahwa *communication cost* mengalami penurunan seiring dengan kenaikan *local epoch* (E), namun disertai dengan kenaikan *computational load*. Pemodelan tren korelasi menggunakan regresi linear menghasilkan R^2 sebesar 0.57 untuk *communication cost* dengan *slope* negatif dan 0,76 untuk *computational load* dengan *slope* positif, menunjukkan adanya *trade-off* antara kedua metrik. Terakhir, titik optimal antara *communication cost* dan *computational load* terletak pada $E = 20$ yang memiliki skor akhir WSM minimum, yaitu 0,2475.

Kata Kunci: *federated learning*, efisiensi komunikasi, FedAvg, IID, penyakit jantung

Heart disease classification using Federated Learning (FL) offers a solution to enhance privacy in the medical field. Unfortunately, FL implementation is constrained by high communication costs that must be optimized while considering a low computational load. This research proposes an FL architecture combining Federated Averaging (FedAvg) with a Multi-Layer Perceptron (MLP) model for heart disease classification using the UCI Heart Disease dataset. This architecture is trained until the accuracy approaches the centralized training scenario as a threshold point across varying local epochs (E). Two metrics are considered in this study: communication cost and computational load, which are modeled using linear regression. Furthermore, the optimal point between both metrics is determined via the Weighted Sum Model (WSM)

method. Test results show that communication cost decreases as the local epoch (E) increases, but is accompanied by an increase in computational load. Correlation trend modeling using linear regression yields an R^2 of 0.57 for communication cost with a negative slope and 0.76 for computational load with a positive slope, indicating a trade-off between both metrics. Finally, the optimal point between communication cost and computational load lies at $E = 20$, which has a minimum final WSM score of 0.2475.

Keywords: *federated learning*, communication efficiency, FedAvg, IID, heart disease

1. Pendahuluan Latar Belakang

Dalam dunia kesehatan, andil teknologi sudah tidak bisa dipungkiri pertumbuhannya. Terlebih lagi sudah banyak peneliti yang mendalami terkait hal tersebut, salah satunya adalah pemanfaatan teknologi *Artificial Intelligence* (AI). Hal ini penting untuk meningkatkan efisiensi dalam melakukan deteksi dan klasifikasi diagnosis penyakit, termasuk penyakit jantung [1]. Penggunaan AI dalam melakukan deteksi penyakit jantung sudah dilakukan oleh [2] dengan akurasi yang mendekati sempurna.

Sayangnya, perkembangan teknologi AI juga diikuti oleh hambatan yang mengancam, yaitu privasi data. Data seperti rekam medis yang bersifat rahasia sangat dikhawatirkan akan keamanannya dari waktu ke waktu [3]. Terlebih lagi jika data-data tersebut dikumpulkan di server terpusat. Hal ini diperkuat dengan regulasi perlindungan privasi data medis berskala internasional untuk AI yang semakin diperketat [4]. Jika terjadi hal-hal yang tidak diinginkan, bukan hal yang mustahil bagi pasien untuk menuntut pihak terkait seperti instansi kesehatan dan para peneliti.

Untuk menjawab keresahan tersebut, hadirilah sebuah pendekatan machine learning yang disebut *Federated Learning* (FL). FL adalah sebuah paradigma yang memungkinkan pembelajaran sebuah model tidak dilakukan secara terpusat. Alih-alih memusatkan semua data ke server terpusat, FL melibatkan beberapa model lokal (klien) yang melakukan pelatihan secara individu [5]. Setelah melakukan pelatihan, setiap model akan memusatkan parameter berupa bobot model

untuk diagregasi secara terpusat. Agregasi yang dilakukan bisa dilakukan dengan menggunakan algoritma *Federated Averaging* (FedAvg) yang berfungsi untuk menghitung rata-rata dari semua bobot kirima klien [6].

Salah satu tantangan dalam mengimplementasikan paradigma FL menggunakan FedAvg adalah tingginya *communication cost*. *Communication cost* bisa diartikan sebagai beban data yang terlibat selama proses pelatihan model. Terdapat beberapa penelitian yang membahas terkait usaha untuk menekan *communication cost*, seperti [6] yang berfokus pada peningkatan *local epoch* sebagai salah satu *hyperparmater* pada komputasi lokal. Namun, kenyataan yang terdapat pada perangkat medis terkini menunjukkan adanya keterbatasan dari sisi kapasitas komputasi, seperti yang disebutkan dalam [7] dan [8]. Hal itu dapat melahirkan risiko lain kegagalan perangkat dan kehabisan sumber daya jika penghematan *communication cost* dilakukan secara tidak terkendali—seperti menaikkan *local epoch* terlalu tinggi. Dengan demikian, penggunaan metrik *communication cost* perlu diseimbangkan dengan metrik *computational load* melalui metode terkait.

Berdasarkan pemaparan di atas, akan dilakukan sebuah analisis *trade-off* antara metrik *communication cost* dan *computational load* sebagai implikasi dari variasi parameter *local epoch*. Selanjutnya, pola korelasi tren dari kedua metrik tersebut akan dimodelkan melalui regresi linear. Sebagai komponen pendukung, akan diimplementasikan metode *Weighted Sum Model* (WSM) untuk menentukan titik optimal dari kedua metrik yang sudah disebutkan sebelumnya. Studi kasus yang digunakan adalah klasifikasi penyakit jantung menggunakan FedAvg dengan skenario data *Independent and Identically Distributed* (IID) yang konvergen terhadap pelatihan terpusat.

Topik dan Batasannya

Topik penelitian ini difokuskan pada pemodelan pengaruh variasi pengaturan *local epoch* terhadap metrik *communication cost* dan *computational load* pada algoritma FedAvg. Secara teknis, input sistem adalah dataset rekam medis penyakit jantung terdistribusi, dan keluarannya adalah model klasifikasi serta penentuan titik optimal arsitektur yang paling efisien. Adapun penelitian ini memiliki beberapa batasan, yaitu menggunakan FedAvg sebagai algoritma FL terhadap dataset UCI Heart Disease yang terdistribusi secara IID. Selain itu, arsitektur yang dibangun menggunakan *Multi-Layer Perceptron* (MLP) sebagai model AI dengan melibatkan regresi linear dalam menganalisis tren dan WSM sebagai metode penentu titik *trade-off*.

Tujuan

Penelitian ini bertujuan untuk memodelkan pengaruh variasi pengaturan *local epoch* terhadap metrik *communication cost* dan *computational load*

pada FedAvg menggunakan metode regresi linear. Selain itu, penelitian ini bertujuan mengimplementasikan metode WSM untuk menentukan titik optimal antara kedua metrik tersebut guna mendapatkan konfigurasi arsitektur yang seimbang.

Tabel 1. Tabel Keterkaitan Antara Tujuan, Pengujian dan Kesimpulan

No.	Tujuan	Pengujian	Kesimpulan
1	Memodelkan pengaruh variasi pengaturan <i>local epoch</i> terhadap metrik <i>communication cost</i> dan <i>computational load</i> pada FedAvg menggunakan metode regresi linear.	Evaluasi	Pemodelan regresi linear membuktikan adanya <i>trade-off</i> ; kenaikan <i>local epoch</i> menghemat <i>communication cost</i> (<i>slope</i> negatif, $R^2 = 0,57$) namun memicu lonjakan <i>computational load</i> (<i>slope</i> positif, $R^2 = 0,76$).
2	Mengimplementasikan metode mendapatkan konfigurasi arsitektur yang seimbang.	Evaluasi menggunakan Min-Max terhadap hasil kedua metrik.	Metode WSM berhasil menentukan konfigurasi <i>local epoch</i> (E) = 20 sebagai titik arsitektur paling optimal dan seimbang, dengan perolehan skor akhir minimum sebesar 0,2475.

Organisasi Tulisan

Jurnal ini disusun dalam beberapa bagian sistematis. Setelah pendahuluan, Bagian 2 (Studi Terkait) memaparkan landasan teori dan penelitian terdahulu mengenai FL dan penyakit jantung. Bagian 3 (Sistem yang Dibangun) menjelaskan alur pra-pemrosesan data hingga desain arsitektur model. Bagian 4 (Evaluasi) menyajikan hasil eksperimen, pemodelan regresi, dan analisis WSM. Jurnal ini diakhiri dengan Bagian 5 (Kesimpulan) yang merangkum temuan utama dan saran pengembangan.

2. Studi Terkait

Penelitian yang dilakukan oleh Maulani *et al.* berfokus pada teknik optimasi *hyperparameter* menggunakan *grid search* dan *random search* pada model hibrida CNN-LSTM [2]. Penelitian lainnya oleh Herlambang memanfaatkan algoritma *Deep Neural Network* (DNN) dalam metode *federated learning* (FL), yang menunjukkan bahwa FedAvg memiliki performa lebih unggul pada skenario data IID [5]. Fondasi pendekatan FL sendiri didasarkan pada penelitian McMahan *et al.* sebagai upaya menangani agregasi data terdesentralisasi, di mana algoritma FedAvg terbukti lebih efisien dibandingkan FedSGD dalam hal biaya komunikasi.

Penelitian berikutnya oleh Narmadha *et al.* membahas penerapan *Federated Efficient Learning* (FedEff) untuk memungkinkan pemilihan *local epoch* optimal pada klien yang heterogen [9]. Pemanfaatan *machine learning* untuk klasifikasi penyakit jantung juga ditunjukkan pada penelitian [10] dengan melibatkan dataset *UCI Heart Disease* menggunakan model *Logistic Regression* (LR) dan *Neural Network* (NN). Dalam penelitian ini, digunakan algoritma *Multi Layer Perceptron* (MLP) yang memiliki keunggulan dalam melakukan aproksimasi fungsi multivariat kontinu [11]. Tabel 2 menjabarkan perbandingan studi terkait dari sisi metode dan hasil penelitian, serta gap penelitian yang diperoleh.

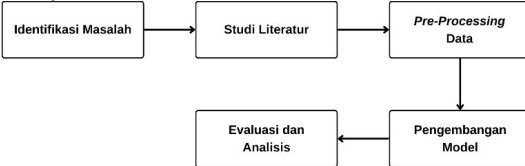
Tabel 2 Tabel Studi Terkait

No.	Referensi	Metode/Algoritma	Hasil Penelitian	Gap Penelitian
1	Maulani <i>et al.</i> [2]	Menggunakan gabungan antara model CNN dan LSTM dengan <i>grid search</i> dan <i>random search</i> .	Berhasil menunjukkan keunggulan <i>grid search</i> dibandingkan <i>random search</i> dalam optimasi <i>hyperparameter</i> .	Belum mengimplementasikan
2	Herlambang <i>et al.</i> [5]	Menggunakan algoritma <i>Deep Neural Network</i> (DNN) yang diterapkan dalam metode <i>federated learning</i> , yaitu FedAvg dan FedProx pada skenario data IID dan Non-IID.	Metode FedAvg memiliki performa yang lebih unggul dari FedProx pada skenario data IID dengan pemilihan fitur yang tepat.	Belum melakukan analisis terhadap <i>communication cost</i> dan <i>computational load</i> pada FL.

3	McMahan <i>et al.</i> [6]	Menggunakan algoritma FedAvg atau <i>Federated Averaging</i> (dibandingkan Non-IID).	Belum mengaitkan metrik <i>communication</i> dengan <i>computational load</i> .
4	Narmadha <i>et al.</i> [9]	Membandingkan efisiensi seperti <i>local epoch</i> secara acak.	FedEff mengurangi <i>epoch</i> acak. Belum mengaitkan metrik <i>communication</i> dengan <i>computational load</i> .

3. Sistem yang Dibangun

Penelitian ini tersusun atas alur sistematis yang meliputi identifikasi masalah, studi literatur, *pre-processing* data, pengembangan model, hingga evaluasi dan analisis seperti yang ditunjukkan oleh Gambar 1.



Gambar 1 Alur Penelitian

Pre-Processing Data

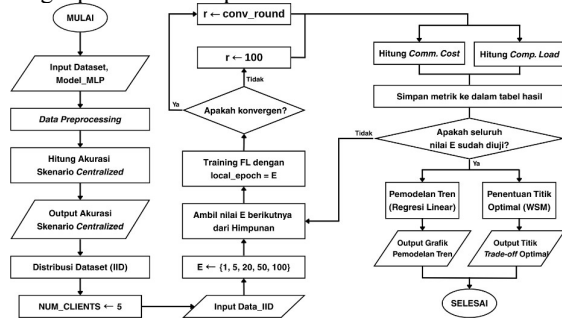
Dataset yang digunakan berjudul *UCI Heart Disease* yang dianotasi secara resmi oleh Janusi *et al.* [12], yang terdiri dari 920 baris data penderita penyakit jantung. Tahapan *pre-processing* meliputi pembersihan kolom yang tidak relevan ('id' dan 'origin'), penyederhanaan target biner, pengodean fitur kategorikal menggunakan one-hot encoding, pengisian nilai hilang dengan median, pemisahan set latihan dan pengujian (80:20), serta standardisasi fitur numerik guna menyamakan skala antarfitur dan menghilangkan dominasi fitur yang memiliki rentang sangat besar [13].

Pengembangan Model

Model MLP yang dibuat terdiri atas 1 *input layer* dengan 13 neuron, 2 *hidden layer* (64 dan 32 neuron), dan 1 *output layer*. Pemilihan jumlah neuron serupa dilakukan dalam penelitian [14] yang turut menggunakan MLP pada studi kasus klasifikasi penyakit jantung. *Hidden layer* menggunakan fungsi aktivasi ReLU, sedangkan output layer memiliki 1 neuron dengan fungsi aktivasi sigmoid untuk memudahkan penentuan probabilitas dalam klasifikasi biner [15]. Skenario pelatihan kedua melibatkan algoritma FedAvg yang bekerja dengan cara mewajibkan setiap klien melakukan pelatihan lokal (*epoch*) lebih dari sekali sebelum mengagregasi bobot model secara terpusat [6].

Pengujian model akan melibatkan 2 skenario pengujian berbeda, namun dengan konfigurasi

model MLP yang sama. Skenario pertama, yaitu skenario terpusat (*centralized*) melibatkan model MLP dengan 4 layer tanpa partisi dataset. Skenario ini akan menghasilkan akurasi yang akan digunakan sebagai baseline titik konvergen (*threshold*) akurasi FL IID dalam mengukur *communication cost* dan *computational load*. Akurasi skenario terpusat diperoleh melalui metode *model checkpoint* yang bekerja dengan cara menyimpan bobot model saat mencapai nilai akurasi validasi tertinggi dari keseluruhan *epoch* [15]. Adapun skenario kedua, yaitu skenario FL IID melibatkan 5 klien dengan distribusi data IID. Variabel *local epoch* (E) diatur bervariasi (1, 5, 20, 50, 100) guna melihat pengaruhnya lebih jauh terhadap metrik evaluasi. Skenario pengujian secara lengkap bisa diamati pada Gambar 2.



Gambar 2 Skenario Pengujian

Evaluasi dan Analisis

Proses evaluasi melibatkan 2 metrik evaluasi, yaitu *communication cost* dan *computational load*. Metrik *communication cost* didefinisikan sebagai total ukuran data yang berpindah dari server ke klien—dan juga sebaliknya—selama proses pelatihan data dalam satuan Byte. Perhitungannya dilakukan melalui rumus

$$C_{comm} = 2 \cdot r_{conv} \cdot n \cdot S_m$$

yang merupakan modifikasi dari penelitian [16] dengan n mewakili jumlah klien (NUM_CLIENTS), r_{conv} mewakili jumlah ronde FL yang konvergen terhadap performa akurasi terpusat, dan S_m mewakili ukuran model dalam satuan Byte. Selanjutnya, metrik *computational load* merupakan kumulasi pemrosesan komputasi yang dibutuhkan guna mengeksekusi pelatihan lokal di setiap klien yang terlibat dalam arsitektur FL. Metrik ini dikalkulasi menggunakan rumus

$$C_{comm} = 2 \cdot r_{conv} \cdot n \cdot S_m$$

dengan E mewakili ukuran *local epoch* yang digunakan dan r_{conv} mewakili ronde FL yang konvergen terhadap performa akurasi terpusat.

Kemudian, dua metrik evaluasi tersebut akan dianalisis melalui metode regresi linear dan *Weighted Sum Model* (WSM). Regresi linear merupakan sebuah metode untuk memprediksi suatu respon kuantitatif melalui pendekatan statistik dengan mempertimbangkan satu prediktor X untuk memprediksi dampaknya terhadap sebuah nilai respon Y [17]. Penelitian yang dilakukan akan melibatkan variasi *local epoch* (E) sebagai

komponen prediktor X —variabel bebas—serta metrik *communication cost* dan *computational load* sebagai nilai respon Y yang diuji secara individu—variabel terikat. Model matematis regresi linear yang digunakan adalah

$$Y = \beta_0 + \beta_1 X + \epsilon \quad (3)$$

dengan Y mewakili hasil metrik evaluasi terikat yang dihitung secara terpisah, X mewakili variabel bebas, β_0 mewakili intersep (nilai perkiraan Y saat $X = 0$), β_1 mewakili koefisien kemiringan atau *slope* untuk mengukur kekuatan dan arah korelasi tren, dan ϵ mewakili galat acak yang tidak bisa dijelaskan oleh model. Adapun metode WSM—salah satu bagian dari kerangka Multi-Objective Optimization (MOO)—merupakan sebuah metode yang memanfaatkan bobot sebagai ukuran yang memanfaatkan tingkat prioritas dari setiap fungsi objektif, relatif satu sama lain [18]. Dengan menggunakan rumus umum WSM yang dicetuskan dalam [18] dengan disertai beberapa penyesuaian, diperoleh formula

$$S_{total} = \alpha \cdot \hat{C}_{comm} + (1 - \alpha) \cdot \hat{C}_{comp}$$

dengan α merupakan parameter preferensi, $\hat{C}_{comm}$ mewakili biaya komunikasi ternormalisasi, dan $\hat{C}_{comp}$ menunjukkan beban komputasi ternormalisasi. Besaran α yang akan digunakan adalah 0,5 karena optimasi menekankan pada keseimbangan preferensi antara kedua fungsi objektif.

4. Evaluasi

Hasil Skenario Terpusat (*Centralized Baseline*)

Pelatihan terpusat menghasilkan akurasi final sebesar 84,78% yang nantinya akan menjadi baseline untuk pelatihan FL menggunakan FedAvg. Target konvergensi FL diambil sebesar 97%⁽¹⁾ dari baseline tersebut, sehingga diperoleh *threshold* sebesar 82,24%.docx (source: 2164, 2165)]. Penetapan nilai tersebut disebabkan adanya distribusi dataset yang independen pada setiap klien, sehingga membuat performa FL belum optimal dalam menyamai performa terpusat [19].

Perilaku Konvergensi *Federated Learning*

Skenario berikutnya menerapkan FL menggunakan FedAvg dengan 5 variasi *local epoch*. Dengan *threshold* 82,24%,⁽²⁾ hasil menunjukkan bahwa peningkatan *local epoch* mempercepat kecepatan konvergensi global: $E = 5$ (ronde 29), $E = 20$ (ronde 8), $E = 50$ (ronde 4), dan $E = 100$ (ronde 2). *Local epoch* yang tinggi membuat pembaruan bobot model menjadi lebih berkualitas sehingga server global memerlukan ronde komunikasi lebih rendah untuk mencapai *threshold* setelah agregasi. Gambar 3 menunjukkan perilaku konvergensi dengan lebih jelas.

Analisis Metrik Evaluasi dan Pemodelan Tren

Pemodelan tren korelasi menggunakan regresi linear terhadap metrik evaluasi menghasilkan dua

temuan yang berbeda untuk setiap metrik evaluasi. Metrik *communication cost* menunjukkan tren berbanding terbalik dengan persamaan

$$Y = 2,9561 + (-0,0322 \cdot X)$$

Angka *slope* (β_1) sebesar -0,0322 menunjukkan adanya penghematan rata-rata sebesar 0,0322 MB untuk setiap peningkatan 1 *local epoch* (X) dengan R^2 sebesar 0,5675. Adapun metrik *computational load* menunjukkan tren berbanding lurus dengan persamaan

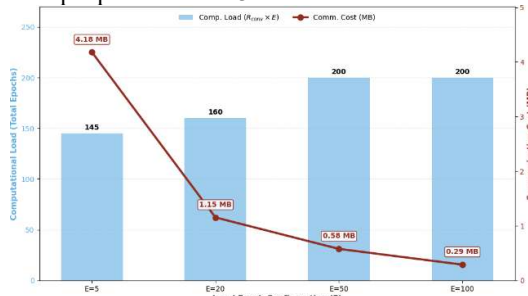
$$Y = 150,6643 + (0,5848 \cdot X)$$

Angka *slope* (β_1) sebesar 0,5848 menunjukkan peningkatan rata-rata sebesar 0,5848 iterasi untuk setiap peningkatan 1 *local epoch* (X) dengan R^2 sebesar 0,7607.

Penentuan Titik Optimal (Trade-off)

Menggunakan WSM

Sebelum masuk ke dalam perhitungan WSM, sangat penting untuk mengamati visualisasi *trade-off* antara *communication cost* dan *computational load* yang terdapat pada Gambar 3.



Gambar 3 Analisis Trade-off Komunikasi-Komputasi

Analisis yang dilakukan tidak mengikutkan skenario $E = 1$ karena skenario tersebut tidak mencapai *threshold* yang sudah ditentukan. Selanjutnya, dilakukan normalisasi terhadap hasil eksperimen menggunakan normalisasi Min-Max. Hasil normalisasi tersebut dikalkulasi lebih lanjut melalui metode WSM dengan $\alpha = 0,5$, karena optimasi menekankan pada keseimbangan preferensi antara kedua fungsi objektif—prioritas setara [20]. Adapun hasil perhitungan lengkap metode WSM terdapat pada Tabel 3.

Tabel 3. Matriks Penentuan Titik Optimal WSM

No.	Local Epoch (E)	C_{comm} (MB)	C_{comp} (Load)	$\hat{C}_{comm}$	$\hat{C}_{comp}$	S_{total}
1	5	4,18 MB	145	1,000	0,000	0,500
2	20	1,15 MB	160	0,222	0,273	0,248
3	50	0,58 MB	200	0,074	1,000	0,537
4	100	0,29 MB	200	0,000	1,000	0,500

Berdasarkan hasil kalkulasi, terlihat bahwa skenario $E = 5$, $E = 50$, dan $E = 100$ tidak menjadi titik optimal karena setiap konfigurasi memiliki nilai normalisasi terburuk. Konfigurasi $E = 5$ memiliki nilai *communication cost* ternormalisasi terburuk, sedangkan konfigurasi $E = 50$ dan $E = 100$ memiliki nilai *computational load* ternormalisasi

terburuk—semuanya bernilai 1,000. Hal ini membuat ketiga skenario tersebut memiliki nilai S_{total} sebesar 0,500 atau lebih. Adapun skenario yang menjadi titik optimal adalah konfigurasi $E = 20$ dengan nilai S_{total} minimum, yaitu sebesar 0,2475. Dengan kata lain, parameter *local epoch* (E) paling optimal untuk arsitektur Federated Learning menggunakan FedAvg pada penelitian ini adalah $E = 20$.

5. Kesimpulan

Berdasarkan penelitian yang sudah dilakukan, diperoleh beberapa kesimpulan. Pertama, peningkatan *local epoch* (E) secara berkala terbukti mampu menurunkan ronde konvergen dari FedAvg yang tentunya juga berdampak kepada penghematan *communication cost* dengan tingkat penurunan yang signifikan pada konfigurasi *local epoch* rendah dan semakin tidak signifikan pada konfigurasi *local epoch* yang lebih tinggi. Kedua, beradaan *trade-off* pada dua metrik evaluasi, yaitu *communication cost* dan *computational load* berhasil dibuktikan melalui pemodelan tren korelasi melalui regresi linear. Hal ini terlihat dari besaran angka *slope* (β_1) negatif pada persamaan regresi linear dari *communication cost* dengan R^2 sebesar 0,5675. Berbeda dengan *slope* pada persamaan regresi dari *computational load* yang besarnya positif dengan R^2 sebesar 0,7607. Ketiga, metode WSM yang diusulkan terbukti mampu menemukan titik *local epoch* (E) optimal dengan mempertimbangkan metrik *communication cost* dan *computational load*, yaitu pada $E = 20$ yang memperoleh skor akhir WSM minimum sebesar 0,2475.

Dengan mempertimbangkan batasan masalah yang menjadi keterbatasan, terdapat dua saran yang bisa diimplementasikan untuk penelitian yang akan datang. Pertama, distribusi data selain IID, yaitu *Non-IID* sebaiknya digunakan untuk penelitian yang akan datang guna melihat respon dan performa dari arsitektur model yang diusulkan. Mengingat data *Non-IID* juga lebih mewakili distribusi data yang ada di dunia nyata. Kedua, arsitektur yang akan ditambahkan sebaiknya melibatkan algoritma lainnya seperti FedProx atau SCAFFOLD untuk menutupi kekurangan FedAvg yang rentan mengalami degradasi performa saat menghadapi heterogenitas klien atau perangkat pada FL (*client drift*).

REFERENSI

- [1] C. W. Tsao *et al.*, "Heart Disease and Stroke Statistics - 2023 Update: A Report from the American Heart Association," 2023. doi: 10.1161/CIR.0000000000001123.
- [2] A. A. Maulani, S. Winarno, J. Zeniarja, R. T. E. Putri, and A. N. Cahyani, "Comparison of Hyperparameter Optimization Techniques in Hybrid CNN-LSTM Model for Heart Disease Classification," *Sinkron*, vol. 9, no. 1, 2024, doi: 10.33395/sinkron.v9i1.13219.
- [3] A. G. Renaudy, "PERLINDUNGAN

- PRIVASI PADA PENGAMBILAN INFORMASI UNTUK BASIS DATA REKAM MEDIS,” Bandung, 2023.
- [4] A. Horst *et al.*, “Federated learning: a privacy-preserving approach to data-centric regulatory cooperation,” *Frontiers in Drug Safety and Regulation*, vol. 5, 2025, doi: 10.3389/fdsfr.2025.1579922.
 - [5] S. W. Herlambang, “FEDERATED LEARNING APPROACHES FOR IOT INTRUSION DETECTION BASED ON FEDAVG AND FEDPROX ON IID AND NON-IID DATA,” Bandung, 2025.
 - [6] H. Brendan McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera y Arcas, “Communication-efficient learning of deep networks from decentralized data,” in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, AISTATS 2017*, 2017.
 - [7] Y. Luo, X. Jin, C. Xia, C. Xu, and Y. Sun, “Communication-Computation Co-Optimized Federated Learning for Efficient Large-Model Embedding Training,” *Mathematics*, vol. 13, no. 23, 2025, doi: 10.3390/math13233871.
 - [8] A. Imteaj, U. Thakker, S. Wang, J. Li, and M. H. Amini, “A Survey on Federated Learning for Resource-Constrained IoT Devices,” *IEEE Internet Things J.*, vol. 9, no. 1, 2022, doi: 10.1109/JIOT.2021.3095077.
 - [9] K. Narmadha and P. Varalakshmi, “FedEff: efficient federated learning with optimal local epochs for heterogeneous clients,” *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-22672-1.
 - [10] J. Premsmith and H. Ketmaneechairat, “A predictive model for heart disease detection using data mining techniques,” *Journal of Advances in Information Technology*, vol. 12, no. 1, 2021, doi: 10.12720/jait.12.1.14-20.
 - [11] K. L. Du, C. S. Leung, W. H. Mow, and M. N. S. Swamy, “Perceptron: Learning, Generalization, Model Selection, Fault Tolerance, and Role in the Deep Learning Era,” 2022. doi: 10.3390/math10244730.
 - [12] A. Janosi, W. Steinbrunn, M. Pfisterer, and R. Detrano, “UCI Machine Learning Repository: Heart Disease Data Set,” *Uci*, 1988.
 - [13] C. M. Bishop and H. Bishop, *Deep Learning: Foundations and Concepts*. 2023. doi: 10.1007/978-3-031-45468-4.
 - [14] E. Saputra, E. R. Susanto, and K. Kunci, “Implementation of Deep Learning with Multilayer Perceptron (MLP) for Heart Disease Prediction Using the SMOTE-ENN Technique,” 2025. [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
 - [15] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola, *Dive into Deep Learning*. Cambridge University Press, 2023.
 - [16] M. Nardi, L. Valerio, and A. Passarella, “Communication Costs Analysis of Unsupervised Federated Learning: an Anomaly Detection Scenario,” in *CEUR Workshop Proceedings*, 2023.
 - [17] G. James, D. Witten, T. Hastie, and R. Tibshirani, “An Introduction to Statistical Learning with Applications in R Second Edition,” *Springer Texts*, 2023.
 - [18] R. T. Marler and J. S. Arora, “The weighted sum method for multi-objective optimization: New insights,” *Structural and Multidisciplinary Optimization*, vol. 41, no. 6, 2010, doi: 10.1007/s00158-009-0460-7.
 - [19] P. Kairouz *et al.*, “Advances and open problems in federated learning,” 2021. doi: 10.1561/22000000083.
 - [20] R. T. Marler and J. S. Arora, “Survey of multi-objective optimization methods for engineering Maximum objective function values,” *Structural and Multidisciplinary Optimization*, vol. 26, no. 6, 2004.