Jurnal Publikasi Teknik Informatika (JUPTI)
Volume.4, Nomor.2 Mei 2025
E-ISSN .: 2808-8972 ; P-ISSN .: 2808-9367, Hal 01-09
DOI: https://doi.org/10.55606/jupti.v4i2.3990
Available online at : https://journalcenter.org/index.php/jupti

Penerapan Federated Learning dalam Keamanan Data Pengguna pada Aplikasi Mobile

Ranto Siswanto [1]*, Muawan Bisri [2]
[1,2] Universitas Indonesia Mandiri, Indonesia
Email : rantoy88@gmail.com [1], muawan_bisri@uimandiri.ac.id [2]*

Abstract: *In the digital era that is full of user data processing, security and privacy are crucial issues, especially in mobile applications. Federated Learning (FL) emerged as an innovative solution in maintaining data confidentiality because the model training process is carried out locally on the user's device without sending raw data to a central server. This study aims to examine the application of FL in improving user data security, evaluate its effectiveness, and analyze the challenges of its implementation in mobile environments. Through a literature study approach and simulated experiments, the results of the study show that FL is able to significantly reduce the risk of data leakage. However, limited device resources and sync issues are major challenges. This research provides important insights into FL's potential in building a more secure mobile app ecosystem.*

Keywords: *data security; distributed machine learning; Federated learning; mobile applications; user privac*

Abstrak: Dalam era digital yang sarat akan pemrosesan data pengguna, keamanan dan privasi menjadi isu krusial terutama pada aplikasi mobile. Federated Learning (FL) muncul sebagai solusi inovatif dalam menjaga kerahasiaan data karena proses pelatihan model dilakukan secara lokal di perangkat pengguna tanpa mengirimkan data mentah ke server pusat. Penelitian ini bertujuan untuk mengkaji penerapan FL dalam meningkatkan keamanan data pengguna, mengevaluasi efektivitasnya, serta menganalisis tantangan implementasinya pada lingkungan mobile. Melalui pendekatan studi literatur dan eksperimen simulatif, hasil penelitian menunjukkan bahwa FL mampu menurunkan risiko kebocoran data secara signifikan. Namun, keterbatasan sumber daya perangkat dan isu sinkronisasi menjadi tantangan utama. Penelitian ini memberikan wawasan penting tentang potensi FL dalam membangun ekosistem aplikasi mobile yang lebih aman.

Kata kunci: aplikasi mobile; Federated learning; keamanan data; machine learning terdistribusi; privasi pengguna

1. PENDAHULUAN

Kemajuan teknologi mobile telah mendorong peningkatan jumlah aplikasi yang mengandalkan data pengguna dalam jumlah besar untuk memberikan layanan yang personal dan efisien. Penggunaan data seperti lokasi, preferensi pengguna, riwayat pencarian, hingga pola perilaku digital kini menjadi fondasi utama dalam membangun sistem yang cerdas dan responsif. Dengan pendekatan berbasis data, aplikasi mampu meningkatkan pengalaman pengguna secara signifikan, seperti dalam hal rekomendasi konten, personalisasi iklan, serta optimalisasi fitur layanan berbasis kebutuhan individu. Namun, seiring meningkatnya kebutuhan akan data, risiko pelanggaran privasi dan kebocoran data menjadi ancaman nyata bagi pengguna (Kairouz et al., 2021).

Dalam beberapa tahun terakhir, publik telah menyaksikan berbagai insiden besar terkait pelanggaran keamanan data, seperti pencurian identitas digital, penyalahgunaan informasi pribadi oleh pihak ketiga, dan kebocoran basis data pengguna dari platform digital populer. Insiden-insiden ini menimbulkan kekhawatiran mendalam terkait sejauh mana kendali

Received: April 15, 2025; Revised: April 30, 2025; Accepted: Mei 04, 2025; Online Available: Mei 07, 2025;

pengguna terhadap data pribadi mereka dapat dipertahankan di tengah era digital yang semakin kompleks. Permasalahan ini memperkuat urgensi untuk mengembangkan metode pengolahan data yang mengedepankan aspek privasi dan keamanan. Pendekatan konvensional yang mengandalkan pengumpulan data terpusat di server utama terbukti rentan terhadap serangan siber dan penyalahgunaan. Oleh karena itu, muncul kebutuhan akan paradigma baru yang mampu menjawab tantangan tersebut, yaitu dengan memungkinkan pengolahan data yang aman tanpa harus mengorbankan privasi individu. Salah satu inovasi yang menjanjikan dalam konteks ini adalah Federated Learning, sebuah pendekatan pembelajaran mesin yang memungkinkan model untuk dilatih secara terdistribusi di perangkat pengguna tanpa perlu memindahkan data mentah ke server pusat. Konsep ini menawarkan solusi yang menarik untuk mengurangi risiko kebocoran data sekaligus tetap mempertahankan kualitas model pembelajaran. Dengan demikian, federated learning tidak hanya menjawab tantangan teknis, tetapi juga menjadi respons strategis terhadap isu etika dan regulasi perlindungan data pengguna yang semakin mengemuka di era digital ini.

2. TINJAUAN LITERATUR

Federated Learning dan Keamanan Data

Federated Learning (FL) merupakan pendekatan pembelajaran mesin terdistribusi yang pertama kali diperkenalkan oleh Google pada tahun 2016 sebagai solusi untuk pelatihan model yang menjaga privasi data pengguna. Pendekatan ini mulai dikenal luas setelah dipublikasikan oleh McMahan et al. (2017), yang menunjukkan bahwa FL memungkinkan pelatihan model dilakukan secara lokal di perangkat pengguna, seperti ponsel pintar, tanpa perlu mengirimkan data mentah ke server pusat. Dalam sistem FL, hanya parameter model atau gradien yang diperoleh dari proses pelatihan lokal yang dikirimkan ke server untuk proses agregasi. Dengan mekanisme ini, data pribadi pengguna tetap berada di perangkat masing-masing, sehingga risiko kebocoran atau penyalahgunaan data dapat diminimalkan secara signifikan. Seiring dengan meningkatnya kesadaran akan pentingnya privasi data, terutama dalam era digital saat ini di mana data menjadi aset berharga, FL muncul sebagai solusi inovatif yang menjembatani kebutuhan antara pelatihan model berkualitas tinggi dan perlindungan privasi. FL tidak hanya menawarkan manfaat dalam menjaga kerahasiaan data, tetapi juga mengurangi kebutuhan akan bandwidth dan penyimpanan pusat, karena data besar tidak lagi perlu ditransfer atau disimpan secara terpusat. Hal ini sangat relevan untuk aplikasi mobile, di mana keterbatasan sumber daya seperti konektivitas, daya baterai, dan penyimpanan menjadi pertimbangan utama.

FL telah digunakan dalam berbagai bidang aplikasi, termasuk pemrosesan bahasa alami (NLP), pengenalan gambar, sistem rekomendasi, dan deteksi anomali. Misalnya, dalam konteks NLP, FL memungkinkan sistem seperti keyboard pintar untuk mempelajari preferensi pengguna tanpa harus mengunggah riwayat ketikan ke server pusat. Dalam pengenalan gambar, FL telah dimanfaatkan dalam pengembangan fitur identifikasi wajah yang dapat dilatih secara kolaboratif di berbagai perangkat. Dalam sistem rekomendasi, FL membantu menciptakan model personalisasi yang lebih akurat tanpa perlu mengorbankan privasi pengguna. Penelitian yang dilakukan oleh Bonawitz et al. (2019) memperkuat temuan awal dari McMahan dan timnya dengan menunjukkan bahwa FL dapat diintegrasikan dengan berbagai teknik keamanan canggih, seperti enkripsi end-to-end dan diferensial privasi. Enkripsi end-to-end memastikan bahwa parameter yang dikirimkan antar perangkat dan server tidak dapat diakses oleh pihak ketiga, sementara diferensial privasi menambahkan noise terkontrol ke dalam parameter model untuk mencegah identifikasi pengguna individu. Kombinasi teknik ini membuat FL menjadi pendekatan yang sangat menjanjikan dalam konteks keamanan data. Meskipun demikian, implementasi FL tidak lepas dari berbagai tantangan teknis dan praktis. Salah satu tantangan utama adalah heterogenitas data dan perangkat.

Data yang tersedia di masing-masing perangkat pengguna dapat sangat bervariasi baik dalam jumlah maupun distribusinya, sehingga proses pelatihan model menjadi kurang stabil. Selain itu, perbedaan dalam kapasitas komputasi dan kondisi jaringan antar perangkat dapat memengaruhi kecepatan dan efisiensi pelatihan model secara keseluruhan. Isu lain yang juga penting adalah keamanan parameter model itu sendiri. Meskipun data tidak dikirimkan, parameter yang dibagikan tetap dapat menjadi sumber informasi yang sensitif jika tidak diamankan dengan baik. Penelitian ini bertujuan untuk menjawab pertanyaan utama: “Bagaimana efektivitas Federated Learning dalam meningkatkan keamanan data pengguna pada aplikasi mobile?” Untuk menjawab pertanyaan tersebut, penelitian ini akan difokuskan pada tiga tujuan utama. Pertama, mengkaji prinsip kerja FL secara mendalam serta relevansinya dalam konteks keamanan data pengguna, khususnya pada aplikasi mobile yang sering menjadi target eksploitasi data. Kedua, mengevaluasi implementasi FL melalui simulasi pada data aplikasi mobile, dengan mengamati aspek kinerja model dan perlindungan privasi. Evaluasi ini akan mencakup perbandingan dengan pendekatan konvensional (centralized learning) dalam hal risiko kebocoran data dan efisiensi pelatihan. Ketiga, mengidentifikasi tantangan implementatif serta potensi pengembangan FL di masa mendatang, termasuk eksplorasi terhadap metode agregasi yang lebih efisien, penggunaan sistem federasi hibrida, dan integrasi dengan teknologi blockchain sebagai lapisan keamanan tambahan. Dengan

pendekatan ini, diharapkan hasil penelitian dapat memberikan kontribusi praktis maupun teoritis terhadap pengembangan sistem pembelajaran mesin yang lebih aman dan andal. Secara praktis, penelitian ini dapat menjadi referensi bagi pengembang aplikasi mobile yang ingin mengimplementasikan teknologi FL untuk meningkatkan kepercayaan pengguna terhadap perlindungan data. Dari sisi teoritis, penelitian ini dapat memperkaya literatur tentang pembelajaran terdistribusi dan membuka ruang diskusi lebih lanjut mengenai kebijakan serta regulasi privasi data di era kecerdasan buatan.

Penerapan FL dalam Aplikasi Mobile

FL telah diterapkan dalam aplikasi keyboard, kesehatan digital, dan pengenalan suara di perangkat mobile (Hard et al., 2018). Tantangan utama dalam penerapan FL pada perangkat mobile adalah keterbatasan daya komputasi, jaringan, dan kebutuhan sinkronisasi model. Beberapa pendekatan seperti penggunaan model ringan dan teknik kompresi telah dikembangkan untuk mengatasi tantangan ini.

3. METODE

Penelitian ini menggunakan pendekatan kuantitatif mixed method yang menggabungkan studi literatur dan eksperimen simulatif untuk menganalisis efektivitas FL dalam meningkatkan keamanan data pengguna pada aplikasi mobile. Sumber literatur diambil dari jurnal bereputasi seperti IEEE, Springer, Elsevier, dan ACM, yang membahas teori dan implementasi FL dalam domain mobile dan keamanan data. Simulasi dilakukan dengan menggunakan framework TensorFlow Federated (TFF) dan dataset Human Activity Recognition Using Smartphones (Anguita et al., 2013), yang umum digunakan dalam penelitian aplikasi mobile. Model FL dibandingkan dengan model sentralisasi konvensional menggunakan parameter evaluasi: Akurasi model, Waktu latih (training time), Keamanan dan risiko kebocoran data (dianalisis dengan metode differential privacy). Alat dan Perangkat : Bahasa pemrograman: Python, Framework: TensorFlow Federated (TFF), Dataset: HAR Dataset, Simulasi pada Google Colab Pro dan perangkat Android virtual, dataset sintetis dibuat berdasarkan pola penggunaan aplikasi nyata, dan model dikembangkan menggunakan pendekatan FL dengan algoritma FedAvg.

Algoritma/Pseudocode

Algoritma 1. Federated Averaging (FedAvg)
Initialize global model weights w0
for each round t = 1, 2, ..., T do:

```
    Select a random subset of clients Ct
    for each client k in Ct do:
        w_k^t = ClientUpdate(k, w_t)
    w_{t+1} = sum_{k=1}^{K} (n_k / n) * w_k^t
return final model w_T
```

Pengujian dan Validasi Model

FL diuji menggunakan metrik akurasi, presisi, dan recall, serta dibandingkan dengan model terpusat.

Pemformatan Komponen Matematika

Rata-rata tertimbang parameter model: $w_{t+1} = \sum_{k=1}^{K} \frac{n_k}{n} w_k^t$

4. HASIL DAN PEMBAHASAN

FL bekerja dengan cara mendistribusikan proses pelatihan ke setiap perangkat. Setelah pelatihan lokal, hanya bobot model yang dikirim ke server untuk dikombinasikan menggunakan algoritma Federated Averaging (FedAvg). Proses ini memungkinkan sistem tetap mendapatkan model yang akurat tanpa harus mengakses data pengguna secara langsung (Bonawitz et al., 2019).

Hasil menunjukkan bahwa FL mampu memberikan akurasi yang kompetitif dengan metode konvensional, namun memiliki keunggulan signifikan dalam aspek keamanan. Risiko kebocoran data secara drastis ditekan karena data tetap berada di perangkat pengguna.

Keterbatasan perangkat: Tidak semua perangkat mobile memiliki daya komputasi memadai untuk melatih model ML. Sinkronisasi komunikasi: FL membutuhkan sinkronisasi yang baik antar perangkat, dan ini menjadi kendala jika konektivitas rendah. Serangan poisoning: FL tetap rentan terhadap manipulasi parameter oleh pengguna jahat jika tidak disertai validasi (Sun et al., 2019).

Penggabungan FL dengan teknik Differential Privacy dan Secure Aggregation dapat meningkatkan keamanan sistem lebih lanjut tanpa mengorbankan kinerja model (Truex et al., 2020).

Gambar dan Tabel

Tabel 1. Ini adalah tabel Perbandingan Akurasi Model.

Metode	Akurasi	Presisi	Recall
FL	87.5%	88.2%	86.9%

Terpusat	89.1%	89.5%	88.7%

Perbandingan Model FL menunjukkan performa yang hampir setara dengan model terpusat, namun dengan keunggulan dalam aspek privasi dan keamanan data pengguna. Waktu komputasi sedikit lebih lama pada FL karena proses distribusi.

Tabel 2. Perbandingan Model FL vs Model Konvensional

Parameter Evaluasi	**Federated Learning**	**Model Konvensional**
Akurasi akhir	89.3%	91.7%
Waktu Pelatihan	65 Menit	45 Menit
Risiko Kebocoran Data	Sangat Rendah	Tinggi

Perbandingan

Model FL menunjukkan performa yang hampir setara dengan model terpusat, namun dengan keunggulan dalam aspek privasi dan keamanan data pengguna. Waktu komputasi sedikit lebih lama pada FL karena proses distribusi. Federated Learning terbukti efektif dalam meningkatkan keamanan data pengguna pada aplikasi mobile tanpa mengorbankan kinerja model. FL merupakan solusi prospektif bagi pengembang aplikasi yang mengedepankan privasi pengguna. Federated Learning memberikan solusi yang inovatif dan aman dalam pengolahan data pengguna pada aplikasi mobile. Dengan menjaga data tetap berada di perangkat pengguna dan hanya mengirimkan parameter model, FL mampu menekan risiko kebocoran data secara signifikan. Hasil simulasi menunjukkan bahwa FL memberikan akurasi yang kompetitif dengan model konvensional, disertai peningkatan privasi dan keamanan. Meskipun masih menghadapi tantangan teknis seperti keterbatasan perangkat dan risiko serangan poisoning, penerapan FL tetap menjanjikan. Kombinasi dengan teknologi keamanan tambahan seperti differential privacy dan secure aggregation dapat menjadi masa depan sistem kecerdasan buatan yang aman dan terpercaya di era digital.

5. KESIMPULAN

Federated Learning terbukti efektif dalam meningkatkan keamanan data pengguna pada aplikasi mobile tanpa mengorbankan kinerja model. FL merupakan solusi prospektif bagi pengembang aplikasi yang mengedepankan privasi pengguna.

Federated Learning memberikan solusi yang inovatif dan aman dalam pengolahan data pengguna pada aplikasi mobile. Dengan menjaga data tetap berada di perangkat pengguna dan

hanya mengirimkan parameter model, FL mampu menekan risiko kebocoran data secara signifikan. Hasil simulasi menunjukkan bahwa FL memberikan akurasi yang kompetitif dengan model konvensional, disertai peningkatan privasi dan keamanan. Meskipun masih menghadapi tantangan teknis seperti keterbatasan perangkat dan risiko serangan poisoning, penerapan FL tetap menjanjikan. Kombinasi dengan teknologi keamanan tambahan seperti differential privacy dan secure aggregation dapat menjadi masa depan sistem kecerdasan buatan yang aman dan terpercaya di era digital.

a) **Kontribusi Penulis:** Penulis bertanggung jawab atas perancangan penelitian, pengumpulan data, analisis, dan penulisan naskah.
b) **Pendanaan:** Penelitian ini tidak menerima pendanaan dari pihak manapun.
c) **Pernyataan Ketersediaan Data:** Data sintesis yang digunakan dalam penelitian ini tersedia atas permintaan kepada penulis.
d) **Ucapan Terima Kasih:** Penulis mengucapkan terima kasih kepada rekan-rekan dosen dan pengembang aplikasi yang telah memberikan masukan terhadap studi ini.
e) Konflik Kepentingan: Penulis menyatakan tidak ada konflik kepentingan terkait penelitian ini.

DAFTAR PUSAKA

Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2013). A Public Domain Dataset for Human Activity Recognition Using Smartphones. *21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN).*

Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., ... & Ramage, D. (2019). Towards Federated Learning at Scale: System Design. *Proceedings of Machine Learning and Systems*, 1, 374–388.

H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), 54, 1273–1282.

Kairouz, P., McMahan, B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Zhao, S. (2021). Advances and Open Problems in Federated Learning. *Foundations and Trends in Machine Learning*, 14(1-2), 1–210.

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 54, 1273–1282.

Sun, J., Wang, X., & Liu, Y. (2019). Can You Really Backdoor Federated Learning?. *arXiv preprint arXiv:1911.07963*.

Truex, S., Baracaldo, N., Anwar, A., Ludwig, H., Zhang, R., & Zhou, Y. (2020). A Hybrid Approach to Privacy-Preserving Federated Learning. *Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security*, 1–11.

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*.

Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., ... & Ramage, D. (2019). Towards federated learning at scale: System design. *arXiv preprint arXiv:1902.01046*.

Hard, A., Rao, K., Mathews, R., Beaufays, F., Augenstein, S., Eichner, H., ... & Ramage, D. (2018). Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604*.

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*.

Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., ... & Ramage, D. (2019). Towards federated learning at scale: System design. *arXiv preprint arXiv:1902.01046*.

Hard, A., Rao, K., Mathews, R., Beaufays, F., Augenstein, S., Eichner, H., ... & Ramage, D. (2018). Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604*.

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Zhao, S. (2021). Advances and open problems in federated learning. *Foundations and Trends in Machine Learning*.

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*.

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*.

Wang, H., Yurochkin, M., Sun, Y., Papailiopoulos, D., & Khazaeni, Y. (2020). Federated learning with matched averaging. *International Conference on Learning Representations*.

Geyer, R. C., Klein, T., & Nabi, M. (2017). Differentially private federated learning: A client level perspective. *arXiv preprint arXiv:1712.07557*.

Zhu, L., Liu, Z., & Han, S. (2019). Deep leakage from gradients. *Advances in Neural Information Processing Systems*.

Mohri, M., Sivek, G., & Suresh, A. T. (2019). Agnostic federated learning. *International Conference on Machine Learning*.

Lin, T., Kong, L., Stich, S. U., & Jaggi, M. (2020). Ensemble distillation for robust model fusion in federated learning. *Advances in Neural Information Processing Systems*.

Smith, V., Chiang, C. K., Sanjabi, M., & Talwalkar, A. (2017). Federated multi-task learning. *Advances in Neural Information Processing Systems*.

Hu, W., Li, Y., & Ding, Y. (2021). A survey of federated learning on edge computing. *Journal of Industrial Information Integration*.

Liu, D., Zhang, Y., Yang, Y., & Xu, D. (2021). Federated learning for smart healthcare: A survey. *IEEE Access*.

Zhao, Y., Li, M., Lai, L., Suda, N., Civin, D., & Chandra, V. (2018). Federated learning with non-IID data. *arXiv preprint arXiv:1806.00582*.

Bagdasaryan, E., Veit, A., Hua, Y., Estrin, D., & Shmatikov, V. (2020). How to backdoor federated learning. *International Conference on Artificial Intelligence and Statistics*.

Nasr, M., Shokri, R., & Houmansadr, A. (2019). Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. *IEEE Symposium on Security and Privacy*.

Yang, K., & Li, X. (2020). Federated learning with adversarial robustness. *IEEE Transactions on Neural Networks and Learning Systems*.

Li, Q., He, B., & Song, D. (2020). Practical one-shot federated learning for cross-silo setting. *arXiv preprint arXiv:2011.13823*.

Chen, M., Yang, Y., Hao, Y., Mao, S., & Hwang, K. (2019). A 5G cognitive system for healthcare. *IEEE Network.*

Kang, J., Xiong, Z., Niyato, D., Zhao, J., & Liang, Y. C. (2020). Incentive design for efficient federated learning in mobile networks: A contract theory approach. *IEEE Internet of Things Journal*.

Sun, Y., Liu, W., Meng, D., & Zhou, D. (2021). Can federated learning save the privacy of your mobile data? *IEEE Transactions on Mobile Computing*.

Li, H., Ota, K., & Dong, M. (2019). Learning IoT in edge: Deep learning for the Internet of Things with edge computing. *IEEE Network.*

Xu, J., & Wang, B. (2020). Privacy-preserving federated learning for edge computing. *IEEE Communications Magazine*.

Ma, C., Liu, J., & Li, H. (2020). Communication-efficient federated learning with adaptive weight aggregation. *IEEE Transactions on Neural Networks and Learning Systems*.