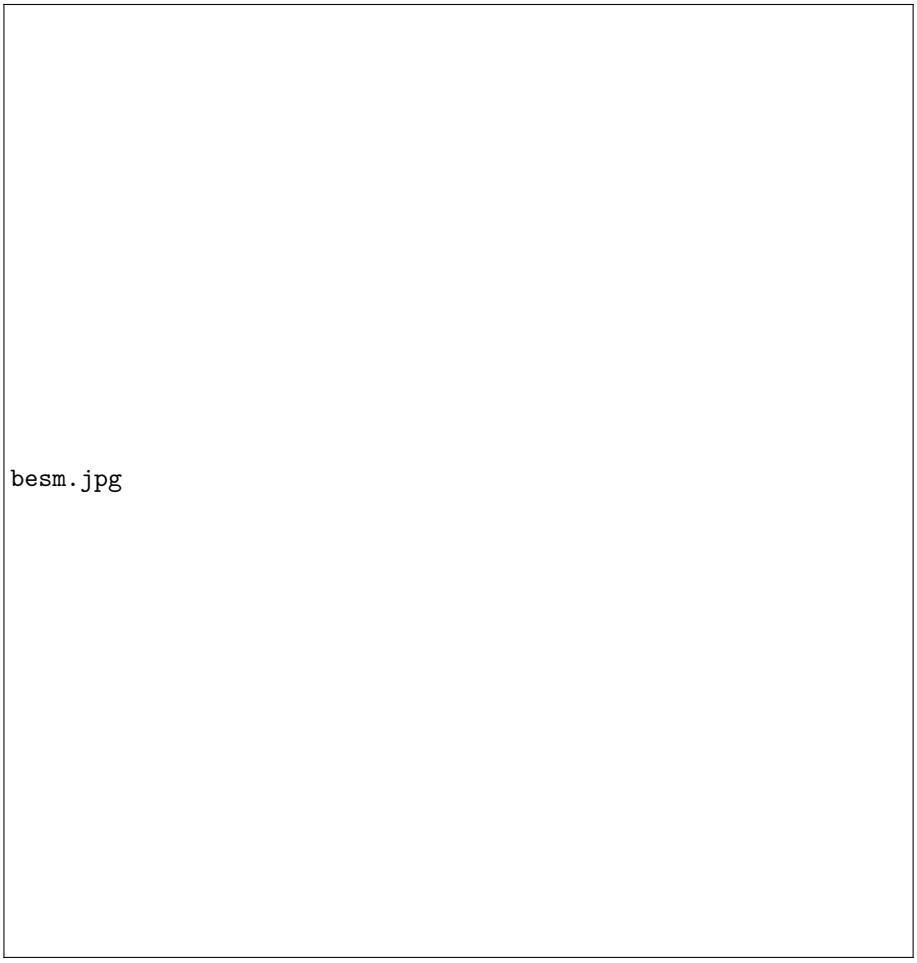
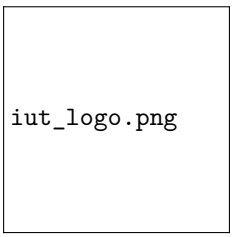

besm.jpg



دانشگاه صنعتی اصفهان
دانشکده مهندسی برق و کامپیوتر

بهبود کارایی الگوریتم یادگیری فدرال برای داده‌های غیرمستقل و غیریکنواخت با در نظر گرفتن میزان شباهت بین شبکه‌های عصبی در دستگاه‌های نهایی

پایان‌نامه کارشناسی ارشد مهندسی کامپیوتر - هوش مصنوعی و رباتیکز

علی بزرگ‌زاد

استادراهنما

دکتر امیر خورسندی

۱۴۰۳

فهرست مطالب

عنوان **صفحه**

چکیده

در عصر حاضر، با پیشرفت سریع فناوری و افزایش تعداد دستگاه‌های متصل به اینترنت، اهمیت برقراری ارتباطات مؤثر و حفاظت از حریم شخصی کاربران بیش از پیش نمایان شده است. این موضوع به توسعه روش‌های توزیع‌شده‌ای مانند یادگیری فدرال انجامیده است. در یادگیری فدرال، داده‌ها به جای ارسال به یک سرور مرکزی، در همان دستگاه‌های نهایی باقی می‌مانند و مدل‌ها به صورت محلی آموزش داده می‌شوند. سپس این مدل‌ها با هم ترکیب می‌شوند تا یک مدل جامع ایجاد شود. این روش نه تنها نیاز به انتقال داده‌ها را کاهش می‌دهد، بلکه به حفظ بهتر حریم شخصی کاربران نیز کمک می‌کند. با این حال، یادگیری فدرال با چالش‌های زیادی روبه‌رو است که یکی از آن‌ها ناهمگنی آماری داده‌ها می‌باشد. به این معنی که داده‌های موجود در دستگاه‌های مختلف می‌توانند بسیار متنوع و متفاوت از یکدیگر باشند. این ناهمگنی باعث می‌شود مدل‌های محلی نتوانند تمامی ویژگی‌های داده‌ها را به‌خوبی یاد بگیرند و در نتیجه، مدل جامع نیز به خوبی همگرا نشود. بنابراین، دستیابی به یک مدل جامع با عملکرد مناسب ممکن است دشوار شود. در این راستا، ارائه روش‌هایی برای مقابله با ناهمگنی آماری از اهمیت بالایی برخوردار است. روش‌های پیشنهادی باید علاوه بر تمرکز بر حل این مشکل، از جنبه‌های محاسباتی، ارتباطی و حفظ حریم شخصی نیز پایداری خود را حفظ کنند. یکی از راهکارهای پیشنهادی برای مقابله با این چالش، جابه‌جایی مدل‌های شبکه عصبی بین کاربران نهایی در طول فرآیند یادگیری است. این کار باعث می‌شود مدل‌های محلی با داده‌های متنوع‌تری مواجه شوند و در نتیجه، مدل جامع به همگرایی بهتری برسد. در روش‌های معمول، جابه‌جایی مدل‌ها به‌صورت تصادفی انجام می‌شود. اما در این پژوهش پیشنهاد شده است که به‌جای روش تصادفی، این جابه‌جایی به‌صورت هوشمند و بر اساس معیارهای شباهت صورت گیرد. به این ترتیب، مدل‌هایی که کمترین شباهت را با هم دارند، جابه‌جا می‌شوند. این رویکرد باعث می‌شود مدل‌ها با داده‌هایی روبه‌رو شوند که کمتر با آن‌ها آشنا هستند و این امر می‌تواند به بهبود همگرایی مدل جامع منجر شود. از جنبه دیگر، این پژوهش به بررسی تأثیر جابه‌جایی مدل‌ها بر حفظ حریم شخصی کاربران پرداخته است. روش‌های معمول جابه‌جایی مدل‌ها به‌طور مستقیم بین کاربران نهایی انجام می‌شوند. اگرچه این روش می‌تواند سربار شبکه را کاهش دهد، اما ممکن است به تضعیف حریم شخصی کاربران منجر شود. در این پژوهش پیشنهاد شده است که سرور مرکزی به عنوان واسطه‌ای در فرآیند جابه‌جایی عمل کند. با این روش، حفظ حریم شخصی کاربران بهتر تضمین می‌شود و پیاده‌سازی تکنیک‌های مختلف این حوزه نیز ساده‌تر خواهد شد. در نهایت، این پژوهش نشان می‌دهد که جابه‌جایی هوشمندانه مدل‌های شبکه عصبی بر اساس معیارهای شباهت، می‌تواند فرآیند همگرایی مدل جامع را تسریع کند. این اثر به‌ویژه در شرایطی که تعداد کاربران زیاد است، مشهودتر خواهد بود.

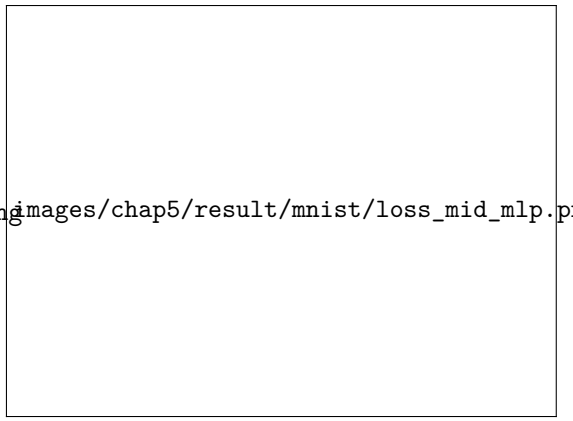
کلمات کلیدی: ۱- یادگیری فدرال، ۲- یادگیری توزیع‌شده، ۳- یادگیری عمیق، ۴- شباهت شبکه عصبی، ۵- ناهمگنی آماری

پیوست اول

بررسی نمودارهای خطا

آ‌ـ۱ مقایسه روش SimFedSwap با روش‌های پایه

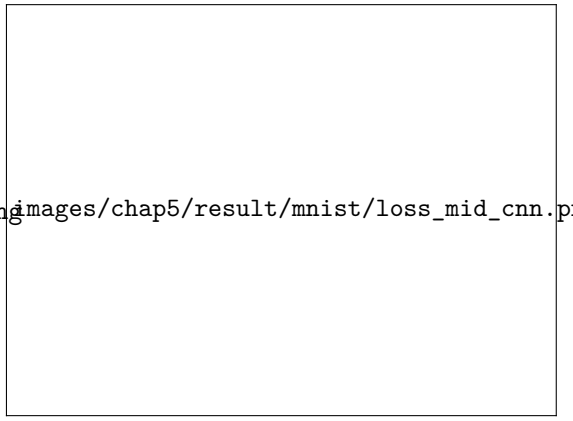
آ‌ـ۱ - ۱ مجموعه داده MNIST



(آ) دید کلی از نتیجه

(ب) بزرگ‌نمایی شده بخش اصلی

شکل آ‌ـ۱: مقایسه منحنی‌های خطا در مجموعه داده MNIST با استفاده از مدل MLP.



(آ) دید کلی از نتیجه

(ب) بزرگ‌نمایی شده بخش اصلی

شکل آ‌ـ۲: مقایسه منحنی‌های خطا در مجموعه داده MNIST با استفاده از مدل CNN.

آ‌ـ۱ - ۲ مجموعه داده CIFAR-10

(آ) دید کلی از نتیجه

(ب) بزرگ‌نمایی شده بخش اصلی

شکل آ‌ـ۳: مقایسه منحنی‌های خطا در مجموعه داده CIFAR-10 با توزیع داده یکنواخت.

(ب) بزرگ‌نمایی شده بخش اصلی (آ) دید کلی از نتیجه

شکل آ‌ـ۴: مقایسه منحنی‌های خطا در مجموعه داده CIFAR-10 با توزیع داده نرمال.

آ‌ـ۱ - ۳ مجموعه داده CINIC-10

(ب) بزرگ‌نمایی شده بخش اصلی (آ) دید کلی از نتیجه

شکل آ‌ـ۵: مقایسه منحنی‌های خطا در مجموعه داده CINIC-10.

آ‌ـ۱ - ۴ مجموعه داده FEMNIST

آ‌ـ۱ - ۴ - ۱ مقایسه نتایج در رویکرد کلاس‌بندی (FEMNISTclass)

(ب) بزرگ‌نمایی شده بخش اصلی (آ) دید کلی از نتیجه

شکل آ‌ـ۶: مقایسه منحنی‌های خطا در مجموعه داده FEMNISTclass.

آ‌ـ۱-۴-۲ مقایسه نتایج در رویکرد نویسندگان (FEMNISTwriter)

(ب) بزرگ‌نمایی شده بخش اصلی (آ) دید کلی از نتیجه

شکل آ‌ـ۷: مقایسه منحنی‌های خطا در یک اجرا بر روی مجموعه داده FEMNISTwriter.

(ب) بزرگ‌نمایی شده بخش اصلی (آ) دید کلی از نتیجه

شکل آ‌ـ۸: مقایسه منحنی‌های خطا در میانگین پنج اجرا بر روی مجموعه داده FEMNISTwriter.

آ‌ـ۲ مقایسه جابه‌جایی حریصانه با جابه‌جایی حداقل شباهت در روش SimFedSwap

(ب) بزرگ‌نمایی شده بخش اصلی (آ) دید کلی از نتیجه

شکل آ‌ـ۹: مقایسه منحنی‌های خطا بین MSS و GS، در مجموعه داده CIFAR-10 با توزیع داده یکنواخت.

(آ) دید کلی از نتیجه (ب) بزرگ‌نمایی شده بخش اصلی

شکل آ‌ـ۱۰: مقایسه منحنی‌های خطا بین MSS و GS، در مجموعه داده FEMNISTwriter.

آ‌ـ۳ تحلیل کاهش تعداد کاربران در هر دور و افزایش تعداد کل دورها در روش SimFedSwap

(آ) دید کلی از نتیجه (ب) بزرگ‌نمایی شده بخش اصلی

شکل آ‌ـ۱۱: مقایسه منحنی‌های خطا در مجموعه داده CINIC-10 با کاهش مشارکت کاربران و افزایش کل دورها.

(ب) بزرگ‌نمایی شده بخش اصلی

(آ) دید کلی از نتیجه

شکل آ‌ـ۱۲: مقایسه منحنی‌های خطا در مجموعه داده FEMNISTclass با کاهش مشارکت کاربران و افزایش کل دورها.

مراجع

[1] Elbir, Ahmet M, Coleri, Sinem, Papazafeiropoulos, Anastasios K, Kourtessis, Pandelis, and Chatzinotas, Symeon. A family of hybrid federated and centralizedlearning architectures in machine learning. *IEEE Transactions on Cognitive Communications and Networking*, 2022.

[2] Zhou, Zhi, Chen, Xu, Li, En, Zeng, Liekang, Luo, Ke, and Zhang, Junshan. Edge intelligence: Paving the last mile of artificial intelligence with edge computing. *Proceedings of the IEEE*, 107(8):1738–1762, 2019.

[3] Ma, Xiaodong, Zhu, Jia, Lin, Zhihao, Chen, Shanxuan, and Qin, Yangjie. A state-of-the-art survey on solving non-iid data in federated learning. *Future Generation Computer Systems*, 135:244–258, 2022.

[4] Smith, Virginia, Chiang, Chao-Kai, Sanjabi, Maziar, and Talwalkar, Ameet S. Federated multi-task learning. *Advances in neural information processing systems*, 30, 2017.

[5] McMahan, Brendan, Ramage Daniel. Federated learning: Collaborative machine learning without centralized training data. https://www.omron.com/global/en/technology/information/dcx, 6 Apr 2017. [Accessed: 18 Apr 2024].

[6] Li, Tian, Sahu, Anit Kumar, Talwalkar, Ameet, and Smith, Virginia. Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine*, 37(3):50–60, 2020.

[7] Talaei, Mahtab. Algorithm development and performance analysis for adaptive differential privacy in federated learning, 21 Aug 2022.

[8] Rieke, Nicola. What is federated learning? https://blogs.nvidia.com/blog/what-is-federated-learning/, 13 Oct 2019. [Accessed: 10 Apr 2024].

[9] Goehner, AIT. Deep learning, welcome to the future! https://www.ait.de/en/deep-learning/. [Accessed: 12 May 2024].

[10] McMahan, Brendan, Moore, Eider, Ramage, Daniel, Hampson, Seth, and y Arcas, Blaise Aguera. Communication-efficient learning of deep networks from decentralized data. in *Artificial intelligence and statistics*, pp. 1273–1282. PMLR, 2017.

[11] Hellström, Henrik, da Silva Jr au2, José Mairton B., Amiri, Mohammad Mohammadi, Chen, Mingzhe, Fodor, Viktoria, Poor, H. Vincent, and Fischione, Carlo. Wireless for machine learning, 2022.

[12] Wang, Hongyi, Sievert, Scott, Liu, Shengchao, Charles, Zachary, Papailiopoulos, Dimitris, and Wright, Stephen. Atomo: Communication-efficient learning via atomic sparsification. *Advances in neural information processing systems*, 31, 2018.

[13] Konečnỳ, Jakub, McMahan, H Brendan, Yu, Felix X, Richtárik, Peter, Suresh, Ananda Theertha, and Bacon, Dave. Federated learning: Strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492*, 2016.

[14] Fang, Chen, Guo, Yuanbo, Hu, Yongjin, Ma, Bowen, Feng, Li, and Yin, Anqi. Privacy-preserving and communication-efficient federated learning in internet of things. *Computers & Security*, 103:102199, 2021.

[15] Konečnỳ, Jakub, McMahan, Brendan, and Ramage, Daniel. Federated optimization: Distributed optimization beyond the datacenter. *arXiv preprint arXiv:1511.03575*, 2015.

[16] Hasan, Jahid. Security and privacy issues of federated learning. *arXiv preprint arXiv:2307.12181*, 2023.

[17] Yin, Xuefei, Zhu, Yanming, and Hu, Jiankun. A comprehensive survey of privacy-preserving federated learning: A taxonomy, review, and future directions. *ACM Computing Surveys (CSUR)*, 54(6):1–36, 2021.

[18] Ioffe, Sergey and Szegedy, Christian. Batch normalization: Accelerating deep network training by reducing internal covariate shift. in *International conference on machine learning*, pp. 448–456. pmlr, 2015.

[19] Li, Tian, Sahu, Anit Kumar, Zaheer, Manzil, Sanjabi, Maziar, Talwalkar, Ameet, and Smith, Virginia. Federated optimization in heterogeneous networks. *Proceedings of Machine learning and systems*, 2:429–450, 2020.

[20] Zhao, Yue, Li, Meng, Lai, Liangzhen, Suda, Naveen, Civin, Damon, and Chandra, Vikas. Federated learning with non-iid data. *arXiv preprint arXiv:1806.00582*, 2018.

[21] Collins, Liam, Hassani, Hamed, Mokhtari, Aryan, and Shakkottai, Sanjay. Exploiting shared representations for personalized federated learning. in *International conference on machine learning*, pp. 2089–2099. PMLR, 2021.

[22] Jeong, Eunjeong, Oh, Seungeun, Kim, Hyesung, Park, Jihong, Bennis, Mehdi, and Kim, Seong-Lyun. Communication-efficient on-device machine learning: Federated distillation and augmentation under non-iid private data. *arXiv preprint arXiv:1811.11479*, 2018.

[23] Taïk, Afaf, Moudoud, Hajar, and Cherkaoui, Soumaya. Data-quality based scheduling for federated edge learning. in *2021 IEEE 46th Conference on Local Computer Networks (LCN)*, pp. 17–23. IEEE, 2021.

[24] Zeng, Yan, Wang, Xin, Yuan, Junfeng, Zhang, Jilin, and Wan, Jian. Local epochs inefficiency caused by device heterogeneity in federated learning. *Wireless Communications & Mobile Computing*, 2022.

[25] Sannara, EK, Portet, François, Lalanda, Philippe, and German, VEGA. A federated learning aggregation algorithm for pervasive computing: Evaluation and comparison. in *2021 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 1–10. IEEE, 2021.

[26] Qin, Yang and Kondo, Masaaki. Mlmg: Multi-local and multi-global model aggregation for federated learning. in *2021 IEEE international conference on pervasive computing and communications workshops and other affiliated events (PerCom Workshops)*, pp. 565–571. IEEE, 2021.

[27] Ma, Qianpiao, Xu, Yang, Xu, Hongli, Jiang, Zhida, Huang, Liusheng, and Huang, He. Fedsa: A semi-asynchronous federated learning mechanism in heterogeneous edge computing. *IEEE Journal on Selected Areas in Communications*, 39(12):3654–3672, 2021.

[28] Li, Li, Duan, Moming, Liu, Duo, Zhang, Yu, Ren, Ao, Chen, Xianzhang, Tan, Yujuan, and Wang, Chengliang. Fedsae: A novel self-adaptive federated learning framework in heterogeneous systems. in *2021 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–10. IEEE, 2021.

[29] Reddi, Sashank, Charles, Zachary, Zaheer, Manzil, Garrett, Zachary, Rush, Keith, Konečnỳ, Jakub, Kumar, Sanjiv, and McMahan, H Brendan. Adaptive federated optimization. *arXiv preprint arXiv:2003.00295*, 2020.

[30] Li, Xiaoli, Liu, Nan, Chen, Chuan, Zheng, Zibin, Li, Huizhong, and Yan, Qiang. Communication-efficient collaborative learning of geo-distributed jointcloud from heterogeneous datasets. in *2020 IEEE international conference on joint cloud computing*, pp. 22–29. IEEE, 2020.

[31] Ghosh, Avishek, Hong, Justin, Yin, Dong, and Ramchandran, Kannan. Robust federated learning in a heterogeneous environment. *arXiv preprint arXiv:1906.06629*, 2019.

[32] Itahara, Sohei, Nishio, Takayuki, Koda, Yusuke, Morikura, Masahiro, and Yamamoto, Koji. Distillation-based semi-supervised federated learning for communication-efficient collaborative training with non-iid private data. *IEEE Transactions on Mobile Computing*, 22(1):191–205, 2021.

[33] Chai, Zheng, Ali, Ahsan, Zawad, Syed, Truex, Stacey, Anwar, Ali, Baracaldo, Nathalie, Zhou, Yi, Ludwig, Heiko, Yan, Feng, and Cheng, Yue. Tifl: A tier-based federated learning system. in *Proceedings of the 29th international symposium on high-performance parallel and distributed computing*, pp. 125–136, 2020.

[34] Jiang, Yihan, Konečnỳ, Jakub, Rush, Keith, and Kannan, Sreeram. Improving federated learning personalization via model agnostic meta learning. *arXiv preprint arXiv:1909.12488*, 2019.

[35] Zhang, Xinwei, Hong, Mingyi, Dhople, Sairaj, Yin, Wotao, and Liu, Yang. Fedpd: A federated learning framework with adaptivity to non-iid data. *IEEE Transactions on Signal Processing*, 69:6055–6070, 2021.

[36] Corinzia, Luca, Beuret, Ami, and Buhmann, Joachim M. Variational federated multi-task learning. *arXiv preprint arXiv:1906.06268*, 2019.

[37] Shoham, Neta, Avidor, Tomer, Keren, Aviv, Israel, Nadav, Benditkis, Daniel, Mor-Yosef, Liron, and Zeitak, Itai. Overcoming forgetting in federated learning on non-iid data. *arXiv preprint arXiv:1910.07796*, 2019.

[38] Chiu, Te-Chuan, Shih, Yuan-Yao, Pang, Ai-Chun, Wang, Chieh-Sheng, Weng, Wei, and Chou, Chun-Ting. Semisupervised distributed learning with non-iid data for aiot service platform. *IEEE Internet of Things Journal*, 7(10):9266–9277, 2020.

[39] Kornblith, Simon, Norouzi, Mohammad, Lee, Honglak, and Hinton, Geoffrey. Similarity of neural network representations revisited. in *International conference on machine learning*, pp. 3519–3529. PMLR, 2019.

[40] Chen, An Mei, Lu, Haw-minn, and Hecht-Nielsen, Robert. On the geometry of feedforward neural network error surfaces. *Neural computation*, 5(6):910–927, 1993.

[41] Orhan, A Emin and Pitkow, Xaq. Skip connections eliminate singularities. *arXiv preprint arXiv:1701.09175*, 2017.

[42] LeCun, Yann, Kanter, Ido, and Solla, Sara. Second order properties of error surfaces: Learning time and generalization. *Advances in neural information processing systems*, 3, 1990.

[43] Gretton, Arthur, Bousquet, Olivier, Smola, Alex, and Schölkopf, Bernhard. Measuring statistical dependence with hilbert-schmidt norms. in *International conference on algorithmic learning theory*, pp. 63–77. Springer, 2005.

[44] Cortes, Corinna, Mohri, Mehryar, and Rostamizadeh, Afshin. Algorithms for learning kernels based on centered alignment. *The Journal of Machine Learning Research*, 13:795–828, 2012.

[45] Cristianini, Nello, Shawe-Taylor, John, Elisseeff, Andre, and Kandola, Jaz. On kernel-target alignment. *Advances in neural information processing systems*, 14, 2001.

[46] Cui, Tianyu, Kumar, Yogesh, Marttinen, Pekka, and Kaski, Samuel. Deconfounded representation similarity for comparison of neural networks. *Advances in Neural Information Processing Systems*, 35:19138–19151, 2022.

[47] LeCun, Yann, Bottou, Léon, Bengio, Yoshua, and Haffner, Patrick. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[48] Holzer, Patrick, Jacob, Tania, and Kavane, Shubham. Dynamically weighted federated k-means. *arXiv preprint arXiv:2310.14858*, 2023.

[49] Krizhevsky, Alex, Hinton, Geoffrey, et al. Learning multiple layers of features from tiny images. 2009.

[50] Carr, Evan Marie. Cifar10 with fast.ai. https://www.evanmarie.com/cifar10-with-fast-ai/, 16 Nov 2022. [Accessed: 4 May 2024].

[51] Darlow, Luke N, Crowley, Elliot J, Antoniou, Antreas, and Storkey, Amos J. Cinic-10 is not imagenet or cifar-10. *arXiv preprint arXiv:1810.03505*, 2018.

[52] Caldas, Sebastian, Duddu, Sai Meher Karthik, Wu, Peter, Li, Tian, Konečnỳ, Jakub, McMahan, H Brendan, Smith, Virginia, and Talwalkar, Ameet. Leaf: A benchmark for federated settings. *arXiv preprint arXiv:1812.01097*, 2018.