Turnitin Originality Report

Processed on: 20-Dec-2024 10:21 PM +08
ID: 2555931952
Word Count: 10505
Submitted: 3

21008537_PD.docx By NG ZU WAYNE
SU.481BSE.202208.FT

Similarity Index	Similarity by Source	
20%	Internet Sources:	14%
	Publications:	15%
	Student Papers:	3%

3% match (Internet from 02-Jul-2024)
https://www.nature.com/articles/s41598-024-63262-x?code=c2d8eb29-3835-446d-849f-9b84dbf602e2&error=cookies_not_supported

2% match (Internet from 27-Jul-2024)
http://arxiv.org/pdf/2407.17999

1% match (Internet from 16-Dec-2022)
https://www.mdpi.com/1424-8220/22/2/450/htm

1% match (Jian Zheng, Cencen Xu, Ziang Zhang, Xiaohua Li. "Electric load forecasting in smart grids using Long-Short-Term-Memory based Recurrent Neural Network", 2017 51st Annual Conference on Information Sciences and Systems (CISS), 2017)
Jian Zheng, Cencen Xu, Ziang Zhang, Xiaohua Li. "Electric load forecasting in smart grids using Long-Short-Term-Memory based Recurrent Neural Network", 2017 51st Annual Conference on Information Sciences and Systems (CISS), 2017

1% match (student papers from 20-Mar-2018)
Submitted to Selçuk Üniversitesi on 2018-03-20

< 1% match (Internet from 16-Sep-2024)
https://www.mdpi.com/1099-4300/26/9/762

< 1% match (Internet from 16-Dec-2024)
https://www.mdpi.com/2071-1050/15/8/6366

< 1% match (Internet from 24-Mar-2022)
https://link.springer.com/article/10.1007/s10845-020-01716-5?code=bde8031a-ae71-44a4-8db9-3bb0ea6899da&error=cookies_not_supported

< 1% match (Internet from 30-Jun-2024)
https://link.springer.com/article/10.1007/s44196-024-00525-5?code=0c26babe-8d44-44c9-b7ab-15424d31c8a3&error=cookies_not_supported

< 1% match (Internet from 16-Sep-2024)
https://link.springer.com/article/10.1007/s11277-024-11372-0?code=f720455c-9fef-404a-a1ae-91483ac970f2&error=cookies_not_supported

< 1% match (Internet from 31-Jan-2023)
https://link.springer.com/article/10.1007/s00521-022-07751-y?code=6e9b439f-98b5-4dcf-acd0-bca2a35b901e&error=cookies_not_supported

< 1% match (Xinyu Wen, Jiacheng Liao, Qingyi Niu, Nachuan Shen, Yingxu Bao. "Deep learning-driven hybrid model for short-term load forecasting and smart grid information management", Scientific Reports, 2024)
Xinyu Wen, Jiacheng Liao, Qingyi Niu, Nachuan Shen, Yingxu Bao. "Deep learning-driven hybrid model for short-term load forecasting and smart grid information management", Scientific Reports, 2024

< 1% match (Fekri, Mohammad Navid. "Deep Learning for Load Forecasting with Smart Meter Data: Online and Federated Learning.", The University of Western Ontario (Canada), 2023)
Fekri, Mohammad Navid. "Deep Learning for Load Forecasting with Smart Meter Data: Online and Federated Learning.", The University of Western Ontario (Canada), 2023

< 1% match (Internet from 14-Sep-2024)
https://ebin.pub/international-conference-on-innovative-computing-and-communications-proceedings-of-icicc-2021-volume-1-advances-in-intelligent-systems-and-computing-1387-981162593x-9789811625930.html

< 1% match (Internet from 06-Dec-2023)
https://ebin.pub/trustworthy-federated-learning-first-international-workshop-fl-2022-held-in-conjunction-with-ijcai-2022-vienna-austria-july-23-2022-revised-selected-papers-3031289951-9783031289958.html

< 1% match (Internet from 24-Dec-2023)
https://ebin.pub/smart-grid-and-sustainable-energy-1774695278-9781774695272.html

< 1% match (Internet from 01-May-2024)
https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2024.1369832/xml/nlm?isPublishedV2=false

< 1% match (Internet from 07-Jul-2024)
https://www.frontiersin.org/journals/built-environment/articles/10.3389/fbuil.2023.1323792/full

< 1% match (Internet from 05-Oct-2022)
https://www.frontiersin.org/articles/10.3389/fenrg.2022.977026/full

< 1% match (Dinh C. Nguyen, Ming Ding, Quoc-Viet Pham, Pubudu N. Pathirana et al. "Federated Learning Meets Blockchain in Edge Computing: Opportunities and Challenges", IEEE Internet of Things Journal, 2021)
Dinh C. Nguyen, Ming Ding, Quoc-Viet Pham, Pubudu N. Pathirana et al. "Federated Learning Meets Blockchain in Edge Computing: Opportunities and Challenges", IEEE Internet of Things Journal, 2021

< 1% match (student papers from 19-Dec-2024)
Submitted to Sunway Education Group on 2024-12-19

< 1% match (Julio Alberto Ramírez-Montañez, Jose de Jesús Rangel-Magdaleno, Marco Antonio Aceves-Fernández, Juan Manuel Ramos-Arreguín. "Modeling of Particulate Pollutants Using a Memory-Based Recurrent Neural Network Implemented on an FPGA", Micromachines, 2023)
Julio Alberto Ramírez-Montañez, Jose de Jesús Rangel-Magdaleno, Marco Antonio Aceves-Fernández, Juan Manuel Ramos-Arreguín. "Modeling of Particulate Pollutants Using a Memory-Based Recurrent Neural Network Implemented on an FPGA", Micromachines, 2023

< 1% match (Internet from 29-Sep-2022)
https://uknowledge.uky.edu/cgi/viewcontent.cgi?article=1126&context=cs_etds

< 1% match (Internet from 13-Dec-2023)
https://jptcp.com/index.php/jptcp/article/download/3649/3587/9620

< 1% match (Mohammad Navid Fekri, Katarina Grolinger, Syed Mir. "Distributed load forecasting using smart meter data: Federated learning with Recurrent Neural Networks", International Journal of Electrical Power & Energy Systems, 2021)
Mohammad Navid Fekri, Katarina Grolinger, Syed Mir. "Distributed load forecasting using smart meter data: Federated learning with Recurrent Neural Networks", International Journal of Electrical Power & Energy Systems, 2021

< 1% match (Sambit Kumar Mishra, Nehal Sampath Kumar, Bhaskar Rao, Brahmendra Brahmendra, Lakshmana Teja. "Role of federated learning in edge computing: A survey", Journal of Autonomous Intelligence, 2023)
Sambit Kumar Mishra, Nehal Sampath Kumar, Bhaskar Rao, Brahmendra Brahmendra, Lakshmana Teja. "Role of federated learning in edge computing: A survey", Journal of Autonomous Intelligence, 2023

< 1% match (Internet from 18-Dec-2021)
https://www.coursehero.com/file/87868748/1c-Introduction-pptx/

< 1% match (G. Dileep. "A survey on smart grid technologies and applications", Renewable Energy, 2020)
G. Dileep. "A survey on smart grid technologies and applications", Renewable Energy, 2020

< 1% match (Mohammad Navid Fekri, Katarina Grolinger, Syed Mir. "Asynchronous adaptive federated learning for distributed load forecasting with smart meter data", International Journal of Electrical Power & Energy Systems, 2023)
Mohammad Navid Fekri, Katarina Grolinger, Syed Mir. "Asynchronous adaptive federated learning for distributed load forecasting with smart meter data", International Journal of Electrical Power & Energy Systems, 2023

< 1% match (student papers from 18-Mar-2019)
Submitted to University of Surrey on 2019-03-18

< 1% match (Internet from 29-Jan-2023)
https://www.researchgate.net/publication/322667510_A_Review_of_Deep_Learning_Methods_Applied_on_Load_Forecasting

< 1% match (Internet from 02-Oct-2022)
https://www.researchgate.net/publication/355046879_Electricity_Forecasting_for_Efficient_Energy_Management_in_Smart_Grid_PhD_Thesis_wi

< 1% match (Mohamed Amine Ferrag, Othmane Friha, Leandros Maglaras, Helge Janicke, Lei Shu. "Federated Deep Learning for Cyber Security in the Internet of Things: Concepts, Applications, and Experimental Analysis", IEEE Access, 2021)
Mohamed Amine Ferrag, Othmane Friha, Leandros Maglaras, Helge Janicke, Lei Shu. "Federated Deep Learning for Cyber Security in the Internet of Things: Concepts, Applications, and Experimental Analysis", IEEE Access, 2021

< 1% match (student papers from 05-Oct-2023)
Submitted to University of Messina on 2023-10-05

< 1% match (Konstantinos Lazaros, Dimitrios E. Koumadorakis, Aristidis G. Vrahatis, Sotiris Kotsiantis. "Federated Learning: Navigating the Landscape of Collaborative Intelligence", Electronics, 2024)
Konstantinos Lazaros, Dimitrios E. Koumadorakis, Aristidis G. Vrahatis, Sotiris Kotsiantis. "Federated Learning: Navigating the Landscape of Collaborative Intelligence", Electronics, 2024

< 1% match (Mamoun Alazab, Suleman Khan, Somayaji Siva Rama Krishnan, Quoc-Viet Pham, M. Praveen Kumar Reddy, Thippa Reddy Gadekallu. "A Multidirectional LSTM Model for Predicting the Stability of a Smart Grid", IEEE Access, 2020)
Mamoun Alazab, Suleman Khan, Somayaji Siva Rama Krishnan, Quoc-Viet Pham, M. Praveen Kumar Reddy, Thippa Reddy Gadekallu. "A Multidirectional LSTM Model for Predicting the Stability of a Smart Grid", IEEE Access, 2020

< 1% match (Syreen Banabilah, Moayad Aloqaily, Eitaa Alsayed, Nida Malik, Yaser Jararweh. "Federated learning review: Fundamentals, enabling technologies, and future applications", Information Processing & Management, 2022)
Syreen Banabilah, Moayad Aloqaily, Eitaa Alsayed, Nida Malik, Yaser Jararweh. "Federated learning review: Fundamentals, enabling technologies, and future applications", Information Processing & Management, 2022

< 1% match (Internet from 11-Jan-2023)
https://pdfcoffee.com/56eye-ball-cursor-movement-using-opencv-pdf-free.html

< 1% match (Filippo Maria Bianchi, Enrico Maiorino, Michael C. Kampffmeyer, Antonello Rizzi, Robert Jenssen. "Recurrent Neural Networks for Short-Term Load Forecasting", Springer Science and Business Media LLC, 2017)
Filippo Maria Bianchi, Enrico Maiorino, Michael C. Kampffmeyer, Antonello Rizzi, Robert Jenssen. "Recurrent Neural Networks for Short-Term Load Forecasting", Springer Science and Business Media LLC, 2017

< 1% match (publications)
Mehdi Ghayoumi. "Generative Adversarial Networks in Practice", CRC Press, 2023

< 1% match (publications)

Rajganesh Nagarajan, Senthilkumar Narayanasamy, Ramkumar Thirunavukarasu, Pethuru Raj. "Intelligent Systems and Sustainable Computational Models - Concepts, Architecture, and Practical Applications", CRC Press, 2024

< 1% match (Theodoros Theodoropoulos, Antonios Makris, Ioannis Kontopoulos, John Violos et al. "Graph neural networks for representing multivariate resource usage: A multiplayer mobile gaming case-study", International Journal of Information Management Data Insights, 2023)
Theodoros Theodoropoulos, Antonios Makris, Ioannis Kontopoulos, John Violos et al. "Graph neural networks for representing multivariate resource usage: A multiplayer mobile gaming case-study", International Journal of Information Management Data Insights, 2023

< 1% match (Vasileios Laitsos, Georgios Vontzos, Apostolos Tsiovoulos, Dimitrios Bargiotas, Lefteri H. Tsoukalas. "Enhanced Sequence-to-Sequence Deep Transfer Learning for Day-Ahead Electricity Load Forecasting", Electronics, 2024)
Vasileios Laitsos, Georgios Vontzos, Apostolos Tsiovoulos, Dimitrios Bargiotas, Lefteri H. Tsoukalas. "Enhanced Sequence-to-Sequence Deep Transfer Learning for Day-Ahead Electricity Load Forecasting", Electronics, 2024

< 1% match (Internet from 19-Dec-2017)
http://btechndcourse.blogspot.com/2015/08/dissertation.html

< 1% match (student papers from 17-May-2024)
Submitted to Coventry University on 2024-05-17

< 1% match (Trang Pham, Truyen Tran, Dinh Phung, Svetha Venkatesh. "Predicting healthcare trajectories from medical records: A deep learning approach", Journal of Biomedical Informatics, 2017)
Trang Pham, Truyen Tran, Dinh Phung, Svetha Venkatesh. "Predicting healthcare trajectories from medical records: A deep learning approach", Journal of Biomedical Informatics, 2017

< 1% match (Haftay Gebreslasie Abreha, Mohammad Hayajneh, Mohamed Adel Serhani. "Federated Learning in Edge Computing: A Systematic Survey", Sensors, 2022)
Haftay Gebreslasie Abreha, Mohammad Hayajneh, Mohamed Adel Serhani. "Federated Learning in Edge Computing: A Systematic Survey", Sensors, 2022

< 1% match (Leo Raju, E Vishal, V Vishwaraj, K M Vimalan. "Application of Machine Learning Algorithms for Short term Load Prediction of Smart grid", 2020 International Conference on Smart Electronics and Communication (ICOSEC), 2020)
Leo Raju, E Vishal, V Vishwaraj, K M Vimalan. "Application of Machine Learning Algorithms for Short term Load Prediction of Smart grid", 2020 International Conference on Smart Electronics and Communication (ICOSEC), 2020

< 1% match (student papers from 21-Apr-2021)
Submitted to University of Glasgow on 2021-04-21

< 1% match (student papers from 04-Sep-2024)
Submitted to University of Hull on 2024-09-04

< 1% match (Internet from 28-Aug-2024)
https://collegepaper1.blogspot.com/2019/09/piezoelectric-energy-home-analysis.html

< 1% match (Internet from 10-Sep-2024)
https://mdpi-res.com/d_attachment/sustainability/sustainability-16-07861/article_deploy/sustainability-16-07861.pdf?version=1725885356

< 1% match (student papers from 27-Apr-2018)
Submitted to Indian Institute of Technology, Bombay on 2018-04-27

< 1% match (student papers from 26-Apr-2017)
Submitted to University of Florida on 2017-04-26

< 1% match (Internet from 07-Sep-2024)
https://www.iww.uni-hannover.de/fileadmin/iww/Institutsmitteilungen/Mit_114_Thesis_Prajna_Kasargodu_Anebagilu_2.pdf

< 1% match (student papers from 06-Jun-2023)
Submitted to Curtin University of Technology on 2023-06-06

< 1% match (Internet from 27-Nov-2024)
https://d197for5662m48.cloudfront.net/documents/publicationstatus/222708/preprint_pdf/6abc1a3dfb4f116df56c3fcac441e79c.pdf

< 1% match (Internet from 08-Oct-2020)
http://orca.cf.ac.uk/135192/13/Wendy%20Harris.pdf

< 1% match (Internet from 03-May-2021)
https://www.journaltocs.ac.uk/index.php?action=browse&journalID=28434&pageb=1&publisherID=450&subAction=pub

< 1% match (Dinh C. Nguyen, Quoc-Viet Pham, Pubudu N. Pathirana, Ming Ding, Aruna Seneviratne, Zihuai Lin, Octavia Dobre, Won-Joo Hwang. "Federated Learning for Smart Healthcare: A Survey", ACM Computing Surveys, 2023)
Dinh C. Nguyen, Quoc-Viet Pham, Pubudu N. Pathirana, Ming Ding, Aruna Seneviratne, Zihuai Lin, Octavia Dobre, Won-Joo Hwang. "Federated Learning for Smart Healthcare: A Survey", ACM Computing Surveys, 2023

< 1% match (Internet from 10-Dec-2024)
https://eprints.gla.ac.uk/342610/1/342610.pdf

< 1% match (Internet from 10-Oct-2024)
https://journalofcloudcomputing.springeropen.com/articles/10.1186/s13677-022-00377-4

< 1% match (Internet from 27-Sep-2022)
https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2778055/no.ntnu%3ainspera%3a83510435%3a83529055.pdf?isAllowed=y&sequence=1

< 1% match (Internet from 13-Nov-2024)
https://sano.science/seminars/application-of-federated-learning-to-medical-data-at-large-scale/

< 1% match (Internet from 11-Nov-2024)
https://wjaets.com/sites/default/files/WJAETS-2024-0384.pdf

< 1% match (Internet from 15-May-2023)
https://www.grin.com/document/1143439?lang=en

< 1% match (Internet from 13-Dec-2024)
https://www.zora.uzh.ch/id/eprint/191299/1/5G%20and%20climate%20protection_University

< 1% match ("Decentralization Technologies", Springer Science and Business Media LLC, 2024)
"Decentralization Technologies", Springer Science and Business Media LLC, 2024

< 1% match (Karki, Bhoj B.. "Assessing Emotive Tasks for User Identification in VR Using Deep Learning", University of Louisiana at Lafayette, 2024)
Karki, Bhoj B.. "Assessing Emotive Tasks for User Identification in VR Using Deep Learning", University of Louisiana at Lafayette, 2024

< 1% match (Malcolm Isaac Fernandez, Yun Ii Go, Mou Ling Dennis Wong, Wolf-Gerrit Früh. "Review of Challenges and Key Enablers in Energy Systems towards Net Zero Target: Renewables, Storage, Buildings, & Grid Technologies.", Heliyon, 2024)
Malcolm Isaac Fernandez, Yun Ii Go, Mou Ling Dennis Wong, Wolf-Gerrit Früh. "Review of Challenges and Key Enablers in Energy Systems towards Net Zero Target: Renewables, Storage, Buildings, & Grid Technologies.", Heliyon, 2024

< 1% match (Neethu Mohan, K.P. Soman, S. Sachin Kumar. "A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model", Applied Energy, 2018)
Neethu Mohan, K.P. Soman, S. Sachin Kumar. "A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model", Applied Energy, 2018

< 1% match (Sumam Francis, Kanimozhi Uma, Marie-Francine Moens. "Understanding the Impact of Label Skewness and Optimization on Federated Learning for Text Classification", Companion Proceedings of the ACM Web Conference 2023, 2023)
Sumam Francis, Kanimozhi Uma, Marie-Francine Moens. "Understanding the Impact of Label Skewness and Optimization on Federated Learning for Text Classification", Companion Proceedings of the ACM Web Conference 2023, 2023

< 1% match (Internet from 16-Sep-2024)
https://dokumen.pub/statistical-data-modeling-and-machine-learning-with-applications-ii-3036582002-9783036582009.html

< 1% match (Internet from 10-Nov-2022)
http://ebot.gmu.edu/bitstream/handle/1920/12673/Chen_gmu_0883E_12633.pdf?isAllowed=y&sequence=1

< 1% match (Internet from 27-May-2022)
https://espace.etsmtl.ca/id/eprint/2992/1/ABDUL_RAHMAN_Sawsan.pdf

< 1% match (Internet from 01-Jan-2022)
https://mobt3ath.com/uplode/books/book-67275.pdf

< 1% match (Internet from 06-Apr-2023)
https://orca.cardiff.ac.uk/id/eprint/158366/1/Alammar%20Thesis%20Final%20D.pdf

< 1% match (Internet from 28-Dec-2022)
https://web.archive.org/web/20220924165453if_/https:/arxiv.org/pdf/2111.09248v3.pdf

< 1% match (Internet from 30-Oct-2024)
https://www.researchsquare.com/article/rs-4909901/v1

< 1% match (Corinne Allaart, Saba Amiri, Henri Bal, Adam Belloum, Leon Gommans, Aart van Halteren, Sander Klous. "Private and Secure Distributed Deep Learning: A Survey", ACM Computing Surveys, 2024)
Corinne Allaart, Saba Amiri, Henri Bal, Adam Belloum, Leon Gommans, Aart van Halteren, Sander Klous. "Private and Secure Distributed Deep Learning: A Survey", ACM Computing Surveys, 2024

< 1% match (Jingbo Zhang, Qiong Wu, Pingyi Fan, Qiang Fan. "A Comprehensive Survey on Joint Resource Allocation Strategies in Federated Edge Learning", Computers, Materials & Continua, 2024)
Jingbo Zhang, Qiong Wu, Pingyi Fan, Qiang Fan. "A Comprehensive Survey on Joint Resource Allocation Strategies in Federated Edge Learning", Computers, Materials & Continua, 2024

< 1% match (publications)
Vandana Mohindru Sood, Yashwant Singh, Bharat Bhargava, Sushil Kumar Narang. "Intelligent Security Solutions for Cyber-Physical Systems", CRC Press, 2024

< 1% match (Internet from 29-Nov-2023)
https://assets.researchsquare.com/files/rs-3648280/v1_covered_488fe5f6-c3ab-4bb8-a3fe-c29f58445061.pdf?c=1701143485

< 1% match (Internet from 23-Nov-2024)
https://internationalpubls.com/index.php/cana/article/download/1463/972/2659

< 1% match (Internet from 14-Apr-2023)
https://ouci.dntb.gov.ua/works/4MZjmxbl/

< 1% match (Internet from 12-Jan-2023)
https://s3-ap-southeast-1.amazonaws.com/ap-st01.ext.exlibrisgroup.com/61RMIT_INST/storage/alma/CC/54/01/20/C9/D9/96/15/48/91/0C/46/1B/EF/6A/D9/Al_Khafaf.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAJN6NPMNGJALPPWAQ%2F20230112%2Fap-southeast-1%2Fs3%2Faws4_request&X-Amz-Date=20230112T212937Z&X-Amz-Expires=119&X-Amz-Signature=de21ade7cf33096a040c9533688be0bd32ff1f13fa848d185931350b365cd990&X-Amz-

SignedHeaders=host&response-content-type=application%2Fpdf

< 1% match (Atif Rizwan, Anam Nawaz Khan, Rashid Ahmad, Qazi Waqas Khan, Do Hyeun Kim. "Personalized hierarchical heterogeneous federated learning for thermal comfort prediction in smart buildings", Engineering Applications of Artificial Intelligence, 2025)
Atif Rizwan, Anam Nawaz Khan, Rashid Ahmad, Qazi Waqas Khan, Do Hyeun Kim. "Personalized hierarchical heterogeneous federated learning for thermal comfort prediction in smart buildings", Engineering Applications of Artificial Intelligence, 2025

< 1% match (Joaquín Delgado Fernández, Sergio Potenciano Menci, Chul Min Lee, Alexander Rieger, Gilbert Fridgen. "Privacy-preserving federated learning for residential short-term load forecasting", Applied Energy, 2022)
Joaquín Delgado Fernández, Sergio Potenciano Menci, Chul Min Lee, Alexander Rieger, Gilbert Fridgen. "Privacy-preserving federated learning for residential short-term load forecasting", Applied Energy, 2022

< 1% match (Yanmin Li, Chao Meng, Jintao Tian, Zhengyang Fang, Huimin Cao. "Data-Driven Customer Online Shopping Behavior Analysis and Personalized Marketing Strategy", Journal of Organizational and End User Computing, 2024)
Yanmin Li, Chao Meng, Jintao Tian, Zhengyang Fang, Huimin Cao. "Data-Driven Customer Online Shopping Behavior Analysis and Personalized Marketing Strategy", Journal of Organizational and End User Computing, 2024

< 1% match (Heba-Allah I. ElAzab, R.A. Swief, Noha H. El-Amary, H.K. Temraz. "Seasonal Electric Vehicle Forecasting Model Based on Machine Learning and Deep Learning Techniques", Energy and AI, 2023)
Heba-Allah I. ElAzab, R.A. Swief, Noha H. El-Amary, H.K. Temraz. "Seasonal Electric Vehicle Forecasting Model Based on Machine Learning and Deep Learning Techniques", Energy and AI, 2023

< 1% match (Minoo Hosseinzadeh, Nathaniel Hudson, Sam Heshmati, Hana Khamfroush. "Communication-Loss Trade-Off in Federated Learning: A Distributed Client Selection Algorithm", 2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC), 2022)
Minoo Hosseinzadeh, Nathaniel Hudson, Sam Heshmati, Hana Khamfroush. "Communication-Loss Trade-Off in Federated Learning: A Distributed Client Selection Algorithm", 2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC), 2022

< 1% match (Mohammad Reza Habibi, Saeed Golestan, Josep M. Guerrero, Juan C. Vasquez. "Deep Learning for Forecasting-Based Applications in Cyber–Physical Microgrids: Recent Advances and Future Directions", Electronics, 2023)
Mohammad Reza Habibi, Saeed Golestan, Josep M. Guerrero, Juan C. Vasquez. "Deep Learning for Forecasting-Based Applications in Cyber–Physical Microgrids: Recent Advances and Future Directions", Electronics, 2023

CAPSTONE PROJECT 1 Planning Document Lightweight Industrial Cohorted Federated Learning (LICFL) for Load Prediction in the Smart Grid by NG ZU WAYNE 21008537 Bachelor of Software Engineering (Hons) Supervisor: Dr Samuel Mofoluwa Ajibade Semester: September 2024 Date: 20 December, 2024 Department of Smart Computing and Cyber Resilience School of Engineering and Technology Sunway University Table of Contents 1.0 Introduction 1.1 Overview 1.2 Background of study 1.3 Problem statement 1.4 Research questions 1.5 Research objectives 1.6 Scope of study 2.0 Literature Review 2.1 Overview 2.2 The Smart Grid 2.3 Load Prediction 2.4 Load Prediction in the Smart Grid 2.5 Federated Learning 2.5.1 Federated Learning Algorithms for Load Prediction 2.5.1.1 Federated Averaging (FedAvg) 2.5.1.2 Federated Stochastic Gradient Descent (FedSGD) 2.5.1.3 FedNorm 2.5.1.4 FLchain 2.5.2 Federated Learning Literature Review Summary Table 2.6 Edge Computing 2.6.1 Edge Federated Learning (EdgeFed) 2.7 Deep Learning 2.7.1 Deep Learning Algorithms for Load Prediction 2.7.1.1 Recurrent Neural Network (RNN) 2.7.1.2 Long Short-Term Memory (LSTM) 2.7.1.3 Gated Recurrent Unit (GRU) 2.7.1.4 Temporal Convolutional Network (TCN) 2.7.2 Deep Learning Literature Review Summary Table 2.8 Conclusion 3.0 Methodology 3.1 Overview 3.2 Research Framework 3.3 Method Implementation 3.4 Datasets 3.5 Lightweight Industrial Cohorted Federated Learning (LICFL) 3.6 Conclusion 4.0 Work Plan and Timeline 4.1 Work Activities for Capstone 1 4.2 Gantt Chart for Capstone 1 4.3 Work Activities for Capstone 2 4.4 Gantt Chart for Capstone 2 5.0 References Chapter 1 1.0 Introduction 1.1 Overview Load prediction is the process of forecasting how much electricity will be needed in the future, and the precise prediction of future energy consumption is simply crucial for advancing the development of the smart energy distribution process (Janjua et al., 2023). In other words, load prediction has played a crucial part in the evolution of the smart grid (Taik & Cherkaoui, 2020). Appropriately, recent advancements in decentralized data processing and machine learning models have introduced innovative approaches to securely and reliably support critical functions in the smart grid, wherein large quantities of energy load data are being utilized on a daily basis to train maching learning models (Hudson et al ., 2021). Edge computing is a modern decentralized computing approach that locates cloud computing services closer to the data sources, near the “edge” of the network. Deep learning is a type of machine learning that applies artificial neural networks to autonomously discover and comprehend complicated trends from huge quantities of data. Edge computing paired with deep learning is a promising technology that has been commonly employed in various applications (Abreha et al., 2022). However, data producers have been required to regularly send their data to external parties in order to train their models. As a result, these applications raise legitimate security concerns, as load profiles expose a lot of restricted information, such as device usage and household occupancy. Sharing such detailed information over these networks makes the information highly susceptible to malicious interception or misuse (Taik & Cherkaoui, 2020). As such, federated learning has recently been proposed as a convincing remedy that addresses the challenge of privacy. This is because federated learning has the ability to train models under distributed clients using a central server while still preserving data localization. Therefore, federated learning has been viewed as a catalyst within the edge computing framework as it allows for collaborative model training and optimization (Abreha et al., 2022). This integration allows for multiple clients to contribute to model improvement without needing to share their raw data, enhancing the effectiveness and privacy of the learning process. Therefore, the main objective of this study is to examine and evaluate the application of Lightweight Industrial Cohorted Federated Learning (LICFL) (Amarlingam et al., 2024) to load prediction. By leveraging LICFL, the study intends to enable efficient and privacy- preserving collaboration among distributed clients, ensuring accurate load prediction while maintaining data security and reducing computational overhead. This approach is particularly suited for the decentralized and heterogeneous nature of the smart grid, where data privacy and computational efficiency are critical. 1.2 Background of study The smart grid has transformed energy distribution by facilitating real time management of energy usage. Load prediction plays an important part in optimizing energy management, and advances in deep learning models have significantly improved prediction accuracy. However, privacy concerns that have been associated with sharing sensitive data in these systems increasingly emphasize the need for federated learning, a privacy-preserving approach. The smart grid is an enhanced version of the traditional electrical grid that utilizes sensing, embedded processing, and digital communication to improve efficiency, automation, and intelligence. Unlike the traditional grid, which is only capable of distributing electric power, the smart grid is able to store data, communicate, and make decisions, allowing for observability, controllability, automation, and integration (Tuballa & Abundo, 2016). Smart meters, a key component, provide consumers with detailed insights into energy usage and costs, facilitate accurate billing, and enable faster outage detection and restoration by utilities. Load prediction is essential for enabling the effective execution of modern power systems. It plays an important part in short term operations and long term strategic planning, especially as the implementation of renewable energy sources like solar power increases. Accurate predictions enable households to enhance self-sufficiency and assist energy planners in making their decisions by evaluating the influence of different variables on consumption. Amid growing market competition, aging infrastructure, and the integration of renewables, load prediction has become indispensable for energy systems planning and optimization (Kuster et al., 2017). Edge computing complements and extends cloud computing by bringing computational responsibilities closer to the data source. Unlike cloud computing, edge nodes handle responsibilities locally, reducing network traffic, minimizing data transmission delays, and improving response times. This decentralized approach not only eases the burden on cloud infrastructure but also enhances security and privacy while addressing challenges such as low latency and real time processing demands (Abreha et al., 2022). Deep learning enables machines to acquire knowledge through experience and comprehend the world through a hierarchy of features, creating complex ideas from simpler ones without the need for explicit human programming. Its application in the smart grid has transformed load prediction by its capability to discover complex patterns and adapt to dynamic energy usage (Biswal et al., 2024). Deep learning models utilize extensive datasets to improve energy usage prediction, load monitoring, and demand response. However, they encounter challenges concerning accuracy and privacy, as effective training requires large, diverse datasets, and smart meter data can potentially expose consumer behavior. Federated learning is a decentralized approach that resolves the problems of communication costs, data privacy, and legal compliance. It enables model training on decentralized clients without the requirement to share raw data, guaranteeing privacy during the training process. A central server averages locally trained model parameters to create and update a global model, which is then redistributed for further local training. This process repeats until the targeted accuracy is attained. Federated learning offers significant advantages over centralized approaches by reducing communication, addressing latency issues, and preserving data privacy, making it suitable for critical applications like industrial automation, mobility systems, real time media, and medical tasks (Reisizadeh et al., 2019). 1.3 Problem statement The smart grid is transforming energy distribution systems. It enables people to monitor, optimize, and predict energy consumption in real time. Load prediction is an important part of this. It allows people to distribute energy efficiently, save money, and keep the grid stable. However, the implementation of predictive models in smart grids often depends on centralized data collection, raising significant concerns regarding privacy and security (Janjua et al., 2023). Sensitive information, such as consumption patterns and user behavior, when centralized, becomes vulnerable to unauthorized access and exploitation. Moreover, data centralization increases the computational burden on central servers, leading to latency issues and inefficiencies. These challenges are further exacerbated by the increasing volume and growing diversity of data being produced by smart grid devices on a daily basis (Powell et al., 2024). Existing approaches to load prediction often struggle to balance accuracy, privacy, and efficiency, limiting their scalability and practicality in the smart grid (Biswal et al., 2024). Therefore, there is a critical need for a machine learning model that enables accurate load prediction while still dealing with the problems of data privacy, distributed processing, and computational overhead. As such, this study aims to face these challenges by evaluating the application of Lightweight Industrial Cohorted Federated Learning (LICFL) to load prediction in the smart grid. 1.4 Research questions A. What are the applications and advantages of federated learning in load prediction? B. How accurate and efficient is the LICFL model for load prediction in the smart grid? C. How does the LICFL model compare to deep learning models in terms of accuracy, efficiency, and performance for load prediction in the smart grid? 1.5 Research objectives A. To identify the applications and advantages of federated learning in load prediction. B. To assess the accuracy and efficiency of the LICFL model for load prediction in the smart grid. C. To compare the accuracy, efficiency, and performance of the LICFL model against deep learning models for load prediction in the smart grid. 1.6 Scope of study This study explores the application of Lightweight Industrial Cohorted Federated Learning (LICFL) (Amarlingam et al., 2024) to load prediction. The scope includes evaluating the LICFL algorithm's ability to enhance efficiency, while preserving privacy. This study concentrates on the challenge of training federated learning models in situations where raw data should never leave local devices due to privacy concerns, while simultaneously reducing the computational overhead that is associated with traditional deep learning models. This study applies the LICFL algorithm to three public datasets: ERCOT Load, AEMO Load, and NYISO. These datasets represent different geographical regions in the wrold with different energy consumption patterns. By applying LICFL to these datasets, this study aims to evaluate its effectiveness in managing varying data distributions in real world scenarios, including temporal and spatial changes in load demand. This study does not cover the design or physical implementation of smart grid infrastructure or edge device hardware. Instead, the focus will be on evaluating the performance of LICFL within smart grid environments. Furthermore, while public datasets are being used, the study will not involve real time data collection or large scale deployment. The primary aim here is to validate the algorithm's functionality and efficiency in load prediction in the smart grid. Chapter 2 2.0 Literature Review 2.1 Overview This literature review provides general introductions to the smart grid, load prediction, federated learning, edge computing, and deep learning. Additionally, the review explores the application of load prediction in the smart grid, the implementation of different federated learning algorithms in load prediction, the application of edge computing in federeated learning, and the implementation of different deep learning algorithms in load prediction. 2.2 The Smart Grid "Smart" suggests intelligence, efficiency, style, or automation, while a grid is a network of electrical conductors that supplies electricity to specific points. Essentially, the smart grid is an advanced version of the traditional grid. The traditional grid can only transmit or distribute electric power. The smart grid can store, communicate, and make decisions, operating in a cooperative, responsive, and organic manner (Tuballa & Abundo, 2016). The smart grid utilizes digital communication technologies to enable the electric grid to be monitored and visualized, manipulated and optimized, capable of adapting and self-healing, and fully compatible with existing systems while accommodating a variety of energy sources (Dileep, 2020). Smart meters track the consumption of electricity, gas, and water. In the smart grid, they provide consumers with insights into their energy use patterns and the cost per kilowatt-hour, leading to more accurate billing and improved pricing information. Additionally, they guarantee quicker outage detection and faster restoration by the utility (Sendin et al., 2014). 2.3 Load Prediction Load prediction, or energy consumption forecasting, is an essential task for the planning, generation, and distribution of energy in contemporary power systems (Khan et al., 2016). It is the process of predicting future energy consumption to ensure a reliable and efficient balance between energy generation and consumption. Short term load prediction is vital for the efficient operation of the energy system, while long term load prediction is a significant part of the strategic planning and development of the system (Fallah et al., 2018). Accurately predicting the immediate electricity consumption of residential consumers has become an area of increasing fascination. The rise in renewable energy adoption, especially solar power, has driven a shift towards more decentralized energy production. Consequently, load prediction has increasingly been perceived as a useful tool for households looking to enhance their self-sufficiency (Peng et al., 2019). Load prediction allows for efficient responses to energy demand. It assists energy planners in comprehending the impact of multiple factors on electricity usage and thus informs decision making (Kuster et al., 2017). Over the past ten years, growing market competition, aging infrastructure and the need to integrate renewable energy have made load prediction increasingly crucial for energy system planning and operation (Hong et al., 2016). 2.4 Load Prediction in the Smart Grid The smart grid plays a significant part in modern electricity systems, utilizing information technology to monitor, manage, and control the energy system in real time. Short term load prediction plays a key part in smart grid management, vital for preserving the stability, efficiency, and reliability of the power network (Raza et al., 2022). It supports safe operations and precise planning while serving as an important tool for risk evaluation and strategic planning in the smart grid (Hasan et al., 2022). Nevertheless, there have been many different uncertainties that affect energy load prediction. These uncertainties have emerged due to the coexistence of patterns and cycles, which is worsened by the substantial randomness and unpredictability (Tang et al., 2021). These make it more difficult to enable precise short term load prediction. Furthermore, the electricity consumption patterns of a household are heavily influenced by the erratic and unpredictable behavior of its residents, making accurate prediction particularly difficult (Kong et al., 2018). 2.5 Federated Learning Federated learning is an approach introduced by Google researchers in 2016, as a convincing answer to the problems of communication, privacy, and authorization (Li et al., 2020). A federated learning approach is a machine learning technique whereby models will be trained on distributed clients that are under centralized control without the need to transmit any raw data. This guarantees data privacy during the training process. An edge or cloud server will repeatedly collect the trained parameters to update a global model, which is then sent back to clients for continued local training (Abreha et al., 2022). The federated learning training process for load prediction generally involves five key steps. The server first selects a machine learning model that will be trained on local data. Second, a random sample of clients will be selected, either randomly or through client selection models like Federated Client Selection (FedCS) (Nishio & Yonetani, 2019). Third, the server multicasts the initial global model to the selected clients, who then download the parameters and train it locally. Fourth, the clients send their updates to the server. Finally, the server averages these updates using aggregation models such as FedAvg (McMahan et al., 2017) to create a new global model without using any raw data. The server administers the training process and distributes the updated global model to clients after every round. The process repeats until the targeted accuracy level is attained. The use of federated learning techniques offers several advantages over traditional centralized machine learning models. Instead of sending raw data, data owners send parameter updates to the server, reducing the volume and size of communication, improving bandwidth efficiency (Reisizadeh et al., 2019). Federated learning addresses latency issues, making it ideal for critical applications. These processes are handled locally on distributed clients, ensuring faster responses (Ang et al., 2020). Additionally, federated learning ensures privacy as raw data remains on local devices, preventing its transmission to a central server. This privacy- preserving technique allows more clients to support the training of a better global model without exposing sensitive information (Nasr et al., 2019). 2.5.1 Federated Learning Algorithms for Load Prediction The main objective in federated learning is to minimize a global loss function that quantifies the overall error of the model across all participating devices. Each device calculates its own local loss function, which measures the error on its local data. The global loss function is computed as a weighted combination of the individual losses that is calculated on each device. The global optimization problem seeks to determine the parameters that minimize the weighted sum of these local losses (Fekri et al., 2021). The weight assigned to each device's contribution is proportional to the size of its local dataset, meaning devices with larger datasets have a greater influence on the overall optimization process. This approach guarantees that the training process accounts for variations and imbalances in the data across devices, effectively integrating insights from all participants to build a robust global model without requiring direct access to their raw data (Fekri et al., 2021). There are many similar yet different federated learning algorithms in related work that employ slightly different approaches to energy load prediction. This literature review takes a brief but concise look into four such federated learning algorithms, namely FedAvg, FedSGD, FedNorm, and FLchain. 2.5.1.1 Federated Averaging (FedAvg) In the FedAvg algorithm, clients send model weights to the server instead of gradients. Upon receiving these weights, the server performs an aggregation step, typically using a weighted average, to create a new global model that represents a consensus across all participating clients. The updated global model is then sent back to clients for the next round. Unlike other methods, FedAvg does not divide training data into batches, which has two results: a substantial reduction in the number of communication rounds, as updates are exchanged only once per epoch, and improved prediction accuracy due to the consistent use of the entire dataset in each round (McMahan et al., 2017). The process repeats until either a fixed number of epochs is reached or a common goal is acheived (Fernández et al., 2022). FedAvg has been shown to significantly improve communication efficiency and reduce training time (Chai et al., 2020). The fundamental principle behind FedAvg is that when all participating clients begin with the same initialization parameters, averaging their weights produces results equivalent to averaging their gradients. This approach ensures that the averaged model performance is not necessarily harmed. However, research has shown that the heterogeneity of data across devices presents challenges, leading to slower convergence rates for FedAvg (Li et al., 2020). Addressing this issue is crucial for improving the model's performance in applications where data distribution varies greatly between clients. 2.5.1.2 Federated Stochastic Gradient Descent (FedSGD) FedSGD employs a decentralized stochastic gradient descent model to collaboratively train a global model within a federated environment. This approach is stochastic in nature because the gradients that are computed on each client are derived from a unbiased sample of the client's local data, rather than the entire dataset. Moreover, the communication between clients and the server presents stochastic elements, as it is often affected by noise, variability in data transmission, and network delays (Husnoo et al., 2023). The randomness in the FedSGD model plays a part in minimizing the risk of overfitting to training data, thereby enhancing the model's capability to generalize and perform accurate predictions on new data. During each round, the stochastic gradient is computed on each device using a randomly selected sample of its local data. The local gradients are then transmitted to the server, where they are averaged using FedSGD to calculate the global gradient. Once the global gradient is obtained, it is used to update the local models, and the process is repeated across successive training rounds (Husnoo et al., 2023). 2.5.1.3 FedNorm FedNorm is a novel asynchronous strategy for load prediction through smart meter data. To deal with challenges such as data staleness, FedNorm calculates the contribution of all the clients by considering the resemblance between the local model weights and the global model weights, as well as the amplitude of the local functions. A global model is then updated according to the contribution of the clients. This approach improves the reduction of training loss in every round, leading to faster convergence. Research has indicated that the model converges speedily under asynchronous and non-IID conditions, outperforming many other synchronous and asynchronous federated learning strategies (Fekri et al., 2023). FedNorm facilitates learning through decentralized smart meter data, even when certain clients are unable to contribute in training as a result of network instability or delays in training. The process begins with Min-Max normalization, applied individually to data from each smart meter, scaling the readings to a range between zero and one. This normalization manages variations in the scale of smart meter data among clients, reduces the influence of larger features, and improves convergence. Following this, the sliding window approach is employed for local data preparation (Fekri et al., 2023). For each smart meter, the first window contains a specified number of consecutive readings, forming the first sample. The window then shifts by a specified number of steps to create the next sample, continuing in this manner to generate subsequent samples. As in regular federated learning, local models train on their local data and send parameters to the server. 2.5.1.4 FLchain Blockchain enables federated learning to be implemented through decentralized data ledgers, removing the requirement of a central server and mitigating the threat of single points of failure. This approach ensures that all update events and user activities are transparently tracked by every entity in the network. Furthermore, transaction logs allow for easy tracing of the origin of model parameter updates during the training process, which is something traditional federated learning systems lack (Majeed & Hong, 2019). The combination of federated learning and blockchain presents an innovative paradigm known as FLchain, which holds the potential to transform intelligent edge networks through its decentralized and secure architecture (Kim & Hong, 2019). In FLchain, all devices function as clients with even rights to update and average the model using a decentralized approach. Essentially, all clients will initialize a model and compute updates. These computed updates are then uploaded to a group of miners as transactions. The miners merge the transactions, which store the clients' updates, into a block after a designated duration. The block is validated by miners under a mining process, and once mined, the block is added to the blockchain and sent out across the network. Clients subsequently download the block and use the block to compute new versions of the global model. This process continues until the global loss function has converged or the required accuracy has been attained. By leveraging blockchain technology, this approach removes the need for a central server, thereby reducing communication cost and achieving improved network scalability. Additionally, blockchain offers enhanced security for federated learning through the application of immutable block ledgers (Nguyen et al., 2021). 2.5.2 Federated Learning Literature Review Summary Table Source Algorithm Strength Limitation Li et al. (2020) FedAvg Enables efficient training by reducing communication while performing multiple local updates on distributed devices. Faces challenges from data heterogeneity, which complicates algorithm design and affects overall performance. Felbab et al. (2019) FedSGD Facilitates efficient training across decentralized data while maintaining performance and reducing communication. Struggles with slow convergence and poor performance on highly imbalanced datasets, affecting overall model accuracy. Fekri et al. (2023) FedNorm Improves performance by adapting client contributions based on local and global model similarities during training. Requires careful management of client contributions to ensure effective model updates and prevent skewed results. Nguyen et al. (2021) FLchain Enhances security and scalability in edge Introduces potential latency and resource networks by enabling management decentralized, challenges due to collaborative training blockchain mining without central server and communication reliance. overhead. Ye et al. (2020) EdgeFed Reduces computational load on clients and minimizes global communication through efficient local updates and aggregation. Requires reliance on edge servers, which may introduce latency and potential points of failure in the network. Amarlingam et al. (2024) LICFL Enhances model performance by clustering similar clients, mitigating issues from heterogeneous data without extra computations. Assumes similarity among clients, which can lead to suboptimal performance in cases of significant data divergence. Wei et al. (2021) Fed-CDP Enhances privacy protection against gradient leakage while maintaining competitive accuracy in federated learning. Incur computational overhead for per- example gradient clipping and sanitization compared to traditional methods. An et al. (2023) FedProx Enhances adaptability to system and statistical heterogeneity, improving convergence and accuracy. Increases complexity of implementation due to additional proximal terms and parameters, complicating the training process. Wang et al. (2023) FTL Improves prediction reliability by transferring learned features from various sources, addressing data sparsity. Complicates model training by requiring careful management of masked load data, which can dilute feature distribution. Husnoo et al. (2023) SignSGD Facilitates efficient communication by transmitting only sign information, minimizing bandwidth usage in distributed learning. Requires careful tuning of hyperparameters to achieve optimal performance, which can complicate implementation. 2.6 Edge Computing Cloud computing offers a vareity of computationally-intense services to users with minimal reliance on hardware resources. However, with the growing demands on the cloud and the rise of delay-sensitive applications, there has been a critical need to look beyond cloud computing (Farhadi et al., 2019). One convincing alternative approach is edge computing, which is also known as fog or micro-cloud computing (Hudson et al., 2021). Edge computing supplements and extends cloud computing. Edge computing integrated with cloud computing offers several benefits over cloud computing alone. The decentralized nature of edge computing allows it to manage numerous computational responsibilities locally, eliminating the need to exchange data with the cloud. This reduces backbone network traffic, minimizes data transmission delays, and enhances response times by hosting services closer to the data source (Abreha et al., 2022). Edge computing is a rising paradigm that aims to solve the problems of response time expectations and data privacy (Ye et al., 2020). It leverages distributed computing to locate computing power and data storage closer to the clients. It minimizes data transmission, enhances service latency, reduces cloud computing strain, and strengthens security and privacy (Abreha et al., 2022). 2.6.1 Edge Federated Learning (EdgeFed) EdgeFed is a federated learning framework designed to improve efficiency in edge computing. By employing an intelligent training technique, it selectively engages edge devices according to resource availability and the relevance of their data. The framework incorporates an optimization approach to optimize the global model while minimizing resource consumption across the network (Mughal et al., 2024). EdgeFed offers several advantages by dividing local model parameter updates between clients and the edge server. Clients concentrate on training the low layers, while the edge server, with its superior resources, handles the more computationally demanding responsibilities. This division allows for lighter and faster training on clients, increasing the approach's practicality. Moreover, the higher bandwidth that is typically available between clients and the edge server compared to that between the edge server and the central server enables EdgeFed to reduce the global communication frequency that is required to achieve satisfactory accuracy. This results in lower global communication cost compared to FedAvg (Ye et al., 2020). The algorithm operates through iterative training between the edge devices, the edge servers, and the central server, followed by global aggregation. Initially, the central server distributes model parameters to the edge servers, which are downloaded by the connected edge devices. Each edge device processes its local data through the lower layers of the model, forwarding the pooling layer's output to its corresponding edge server. The edge servers aggregate these outputs into larger matrices to serve as inputs for subsequent layers. Using gradients and a loss function, the edge servers update the model parameters and return them to the edge devices. After several split training iterations, the edge servers send the updated parameters to the central server, where they are aggregated based on weights. The globally updated parameters are then redistributed to the edge servers for further training (Ye et al., 2020). 2.7 Deep Learning Deep learning enables machines to learn from experience and comprehend the world through a hierarchy of features. As the machine gains knowledge from experience, there is no requirement for a human to specify the knowledge that is required by the machine. The hierarchy of features enables the machine to understand complex concepts by creating them out of simpler concepts; a graph of the hierarchies will be multiple layers deep (Heaton, 2017). The application of deep learning has significantly transformed how load prediction operates within the smart grid. Deep learning, particularly neural networks, have shown an exceptional capability in discovering complicated relationships and patterns within large amounts of data. The ability of deep learning to autonomously extract features and adapt to changing situations makes deep learning very suitable for the dynamic nature of energy consumption. Significant progress in energy usage estimation and prediction has been achieved through deep learning models (Biswal et al., 2024). In the smart grid, huge amounts of energy consumption data are being utilized to train deep learning models. Nevertheless, these applications experience challenges involving security and possess high accuracy requirements. Privacy is a key concern, as the data being collected by smart meters can reveal information about household appliances and thus consumer behavior. Furthermore, deep learning models require large and diverse datasets for adequate training (Taik & Cherkaoui, 2020). 2.7.1 Deep Learning Algorithms for Load Prediction Deep learning has garnered significant attention in recent years for its ability to handle large datasets and identify complex patterns. In load prediction, deep learning methods present innovative approaches for modeling and predicting electricity consumption, effectively addressing many limitations of traditional prediction techniques (Yazici et al., 2022). Deep learning models have the ability to catch complex long term relationships and non-linear trends. Their ability to manage large quantities of data makes them very compatible with applications that involve large data flows. The adaptable nature of deep learning models facilitates seamless integration with other models, enhancing prediction accuracy. Nevertheless, training deep learning models necessitates substantial processing power. Furtheermore, their sophistication makes them prone to the overfitting of training data, potentially resulting in poor generalization to new data unless regularization techniques are effectively applied (Haque & Rahman, 2022). There are many similar yet different deep learning algorithms in related work that employ slightly different approaches to energy load prediction. This literature review takes a brief but concise look into four such deep learning algorithms, namely RNN, LSTM, GRU, and TCN. 2.7.1.1 Recurrent Neural Network (RNN) A Recurrent Neural Network (RNN) is a kind of artificial neural network that is aimed at processing and analyzing sequential data, such as text in language translation responsibilities or time series information in load prediction. The defining feature of an RNN is its recurrent connections, which link the current output to its previous state, enabling the network to maintain an internal memory. This architecture makes RNNs adept at catching temporal relationships and modeling sequential behavior over time. However, training RNNs typically involves backpropagation through time. For lengthy sequences, this method can result in the vanishing gradient problem, where gradients diminish as they propagate backward, leading the network to struggle in retaining older information (Fekri et al., 2020). RNN is a learning machine that repeatedly computes new states by employing transfer functions in previous states and current inputs. These transfer functions consist of an affine transformation that is followed by a nonlinear activation function, with the specific choice of function tailored to the nature of the problem being addressed. RNNs are known to possess the universal approximation property, meaning they can approximate arbitrary nonlinear dynamical systems with any desired level of precision. This capability enables them to build complicated mappings from input sequences to output sequences. The architecture of an RNN plays a significant role in determining how data flows between its neurons, and its intricate architecture is important for attaining an effective system (Bianchi et al., 2017). In prediction, an RNN will be trained on input data to produce a corresponding temporal output. The output can represent any time series that is related to the input, including a time-shifted version of the input itself. While gradient-based methods are the most commonly used for training, other approaches such as derivative-free methods or convex optimization techniques have also been proposed (Bianchi et al., 2017). The training process aims to minimize a loss function that quantifies the error between the network's predicted output and the actual target output. Notably, after sufficient training, RNNs can be executed in a generative mode, replicating temporal patterns similar to those observed in training data (Gregor et al., 2015). 2.7.1.2 Long Short-Term Memory (LSTM)

LSTM is an efficient RNN designed to address the vanishing gradient problem encountered by standard RNNs when processing long term relationships. In a standard RNN, the overall neural network consists of a sequence of repeating modules, each created through a simple hidden network. In contrast, the hidden layers of LSTM possess a more complex design. LSTM presents the concepts of gates and memory cells within each hidden layer. Each memory block is composed of four components: the input gate, the forget gate, the output gate, and the self- connected memory cells (Zheng et al., 2017). The input gate manages the activations that enter the memory cell. The output gate determines when to release the activations into the next layer in the network. The forget gate enables the model to discard irrelevant past information and reset the memory cells. Additionally, multiplicative gates will be employed with precision, allowing the memory cells to store and access data over extended periods. This structure effectively solves the vanishing gradient problem, making LSTM suitable for responsibilities that involve long term relationships (Zheng et al., 2017). However, when the output gate has closed, the gates cannot access information from the memory cell output, leaving LSTM unable to determine the required duration for retaining the memory. To solve this challenge, peephole connections have been integrated into the memory cells. Acting as direct supervisors, peephole connections enable all gates to monitor the cell states (Zheng et al., 2017). 2.7.1.3 Gated Recurrent Unit (GRU) GRU is a type of RNN that addresses the vanishing gradient problem, a regular problem in traditional RNNs whereby gradients diminish during backpropagation, hindering the network's capability to learn from long sequences. GRUs address this issue through gating mechanisms that manage the flow of data within the network. These mechanisms consist of two main gates: the update gate and the reset gate. The update gate manages how much of the previous hidden state is integrated with the current state, effectively balancing past information and new input. Meanwhile, the reset gate decides how much of the previous hidden state is reset before being merged with current input, allowing the GRU to selectively discard irrelevant historical data (Wen et al., 2024). GRU is pivotal in identifying and retaining long term relationships in time series information. By handling sequential input data, its gated units understand to preserve essential data from previous time steps to pass forward to the current time step. This capability to catch and leverage relationships is crucial for achieving precise and reliable short term load prediction (Wen et al., 2024). 2.7.1.4 Temporal Convolutional Network (TCN) The Temporal Convolutional Network (TCN) employs dilated causal convolutions to catch local and global relationships within input sequences. The fundamental principle behind TCN involves its usage of dilated causal convolutions, which expand the receptive fields exponentially as the depth of the network increases. This enables TCN to model long term relationships in input sequences, making it effective for responsibilities involving temporal relationships (Wen et al., 2024). In TCN, the input sequence is handled by a collection of convolutional layers, where every layer uses filters to extract local patterns and features. A distinguishing feature of TCN is its usage of dilated convolutions, which present gaps (dilations) in between filter elements. By progressively increasing the dilation rate across layers, TCN effectively captures relationships over short and long time spans. To maintain causality, padding has been employed in the input sequence, ensuring that the filters access only past and current time steps, thereby preventing future data leakage. This approach enables TCN to have a huge receptive field while keeping the amount of parameters manageable, enhancing computational efficiency (Wen et al., 2024). 2.7.2 Deep Learning Literature Review Summary Table Source Algorithm Strength Limitation Bianchi et al. (2017) RNN Captures temporal dependencies in data, enabling effective modeling of complex, nonlinear relationships for accurate predictions. Experiences difficulties in remembering long sequences due to limited memory capacity and suboptimal parameter training. Kong et al. (2017) LSTM Models sequential dependencies effectively, allowing for better prediction of individual residential energy consumption trends. Struggles with high volatility in individual load patterns, leading to less accurate prediction in certain unpredictable scenarios. Li et al. (2022) GRU Optimizes parameter efficiency, requiring fewer resources for training while achieving competitive prediction results. Can be sensitive to the choice of hyperparameters, which significantly affects its performance and convergence during training. Song et al. (2020) TCN Improves prediction precision by leveraging multi-scale temporal patterns in data through a deep learning architecture. Can suffer from overfitting if not regularized adequately, especially with limited training data. Hosein and Hosein (2017) DNN Captures complex patterns in electrical load data through multiple hidden layers, resulting in superior predictive performance. Requires significantly longer computational times compared to traditional methods, impacting efficiency in dynamic prediction environments. Li et al. (2020) DCN Improves prediction precision by utilizing a densely connected architecture and advanced regularization methods to mitigate overfitting. May require extensive computational resources and time due to its complex architecture and numerous parameters. Tayab et al. (2020) HHO-FNN Improves prediction reliability by minimizing error fluctuations and leveraging advanced optimization techniques for better training. Can be sensitive to initial conditions, potentially affecting the consistency of prediction results. Real et al. (2024) DRL Enhances self- consumption of renewable energy by intelligently managing storage and load based on real-time data. Requires substantial computational resources and extensive data for training, which can complicate implementation. Guo et al. (2024) DBN Enhances predictive accuracy and efficiency in processing high- Struggles with suboptimal initial weight selection, potentially leading to dimensional, complex datasets for improved operational optimization. poor predictive performance and local optima issues. Arvanitidis et al. (2022) MLP Enhances prediction accuracy and convergence speed by integrating clustering techniques with neural network inputs. Struggles with convergence speed and accuracy when not informed by effective data preprocessing techniques like clustering. 2.8 Conclusion This literature review has collated diverse insights into the different federated learning and deep learning algorithms that have been employed in load prediction in related work. As such, the review has appropriately consolidated the foundational knowledge that will be required to evaluate the application of Lightweight Industrial Cohorted Federated Learning (LICFL) to load prediction in the smart grid. Chapter 3 3.0 Methodology 3.1 Overview This study proposes the application of the LICFL algorithm (Amarlingam et al., 2024) to address the problems of privacy and accuracy in load prediction in the smart grid. Essentially, the LICFL algorithm groups similar clients into cohorts and performs federated learning within each cohort. Therefore, the LICFL algorithm incorporates a model parameter based cohorting algorithm. This is because the LICFL algorithm utilizes model parameters to cohort clients, making the approach lightweight and eliminating the requirement for any supplementary communications or computations at the client level for cohorting. Additionally, the LICFL algorithm incorporates an adaptive strategy selection algorithm. The adaptive strategy selection algorithm takes advantage of four different aggregation strategies, namely FedAvg, FedAdaGrad, FedYogi, and FedAdam, through selecting the optimal strategy for each communication round. This approach improves global model convergence and performance. 3.2 Research Framework 3.3 Method Implementation This study aims to contrast the performance of the LICFL model with deep learning models in the task of load prediction. The models will be compared using four accuracy metrics (Accuracy, Recall, F1 Score, AUC), four efficiency metrics (Parameters, Flops, Inference Time, Training Time), and four evaluation metrics (MAE, MAPE, RMSE, MSE). The method implementation is as follows: 1. Data Preparation: Select suitable load datasets that contain historical energy consumption data. Ensure the datasets are of adequate quality and relevance. 2. Feature Engineering: Extract and select features from the datasets that represent key aspects of load prediction. 3. Data Preprocessing: Preprocess the datasets by handling missing values, managing outliers, and applying normalization. Split the datasets into training and testing sets. Ensure the relevance of the testing set. 4. Model Construction: Construct the LICFL algorithm, the model parameter based cohorting algorithm, and the adaptive strategy selection algorithm. 5. Model Training: Configure model hyperparameters. Train the model using the training set. Record the training time. 6. Model Inference: Use the testing set to conduct inference with the trained model. Record the inference time. 7. Comparative Experiment: Select deep learning models as comparative models. Ensure the same datasets have been applied. 8. Results Analysis: Compare the performance of the models across the various metrics. Analyze their strengths and weaknesses. 9. Conclusion: Summarize the models' strengths and weaknesses that have been identified in the results analysis. Discuss the significance of the various metrics in load prediction. Provide recommendations for future improvements based on the findings. 3.4 Datasets The following three datasets have been selected for this study: 1. ERCOT Load dataset: The dataset contains load consumption data that is managed by the Electric Reliability Council of Texas (ERCOT). Notably, the ERCOT region is one of the largest electricity markets in the United States. The dataset offers comprehensive historical records of energy load data, enabling the analysis and modelling of energy consumption trends in the ERCOT region. However, its focus on the ERCOT region limits its applicability to other markets, and the lack of certain supplementary data, such as weather or economic varaibles, may reduce the accuracy of energy prediction models. 2. AEMO Load dataset: The dataset contains load consumption data from the Australian Energy Market Operator (AEMO), capturing energy consumption trends in the Australian energy market. The dataset offers information regarding local energy consumption trends, enabling the development of region-specific energy prediction and management strategies. However, its applicability outside of Australia may be limited, as it may not capture factors unique to other regions. Furthermore, the dataset may also lack certain contextual elements, such as demographic or environmental variables, which may also influence energy consumption patterns. 3. NYISO dataset: The dataset contains load consumption data from the New York Independent System Operator (NYISO), capturing energy consumption trends in the state of New York, USA. The dataset offers a comprehensive view of local energy consumption trends, enabling the development and testing of energy prediction and management strategies specific to the region. However, the dataset may not cover certain variables that may also influence electricity demand, such as local events or specific market fluctuations. 3.5 Lightweight Industrial Cohorted Federated Learning (LICFL) Generally, the weights (or model parameters) of a neural network have been trained as a function of input data (Shalev-Shwartz & Ben-David, 2015). Therefore, the weights capture the input data's properties. Based on this understanding of model parameters, Algorithm 1 introduces a lightweight approach that only evaluates the model weights or parameters. This approach removes the need for any supplementary communications or computations at the client level. The pseudocode of the LICFL algorithm is displayed in Algorithm 1. In the first round, the server begins by sending the initial model to clients (Algorithm 1: Line 6). Clients then update the initial model based on the loss function and their respective local data (Algorithm 1: Lines 30 to 32). Next, clients send the updated model parameters back to the server (Algorithm 1: Line 33). After receiving the set of updated parameters from the clients (Algorithm 1: Line 8), the server executes the model parameter based cohorting algorithm (Algorithm 1: Line 10). The pseudocode of the model parameter based cohorting algorithm is displayed in Algorithm 2. Neural network models are composed of multiple layers. The model parameters for each client are flattened into a single vector for further processing (Algorithm 2: Line 3). The model parameter vectors from the clients are concatenated into a matrix. However, high dimension clustering leads to inaccurate clusters (Pandove et al., 2018). To address this, Algorithm 2 compresses the model parameters into a lower dimension while maintaining the relevant features with high variability (Algorithm 2: Lines 5 to 8). To address the complex relationships within the model parameters, graph based clustering is employed, taking inspiration from spectral clustering. A weighted graph is built using an adjacency matrix that is derived from the transformed matrix (Algorithm 2: Line 9), and the Laplacian matrix is computed from it (Algorithm 2: Line 11). Through eigenvalue decomposition of the Laplacian matrix (Algorithm 2: Line 12), the largest eigenvalues and the corresponding eigenvectors are identified (Algorithm 2: Line 13). The eigenvectors are then normalized to create a new matrix to be employed for clustering (Algorithm 2: Line 14) where the rows stand for the clients and the columns stand for the transformed weights belonging to the clients in the new space. For this implementation, k-means clustering will be applied for the grouping of the columns of the normalized matrix. The output of Algorithm 2 is a set of cohorts (Algorithm 2: Line 15). In subsequent rounds, Algorithm 1 updates the parameters for each cohort to construct personalized models for the clients within it. Initially, the server sends the model parameters for each cohort to all the clients within it (Algorithm 1: Line 19). The clients in the cohort train the model locally and send the updated model parameters to the server. These updates are combined using an aggregation method, resulting in a new set of updated model parameters for the cohort (Algorithm 1: Line 22). The server then redistributes the aggregated model parameters to the clients in the cohort (Algorithm 1: Line 24). Each client updates its local model using these aggregated parameters (Algorithm 1: Line 26). This process repeats for every round across all cohorts. Different aggregation strategies have been observed to perform better at different rounds of communication. To take advantage of this and further enhance global model performance and convergence, an adaptive aggregation strategy selection method is proposed. Algorithm 3 enhances global model convergence within each cohort of federated learning by selecting the optimal aggregation strategy solely based on model parameters. Additional communications or computational resources at the client level are not required. The pseudocode of the proposed adaptive strategy selection algorithm is displayed in Algorithm 3. Common momentum-based global model optimization methods propose different aggregation strategies to address the variations in data among the different clients. However, these methods each depend on a single aggregation strategy for every round, which does not provide the optimal result. Algorithm 3 improves upon this by dynamically selecting the most suitable aggregation strategy (FedAdam, FedAdaGrad, FedYogi, FedAvg) for each round. This adaptability improves model performance and convergence. Algorithm 3 begins by taking as input a set of model parameters from clients at the current round, along with the initial aggregated model parameters or the parameters from the previous round. Momentum parameters are then calculated for the different aggregation strategies (Algorithm 3: Lines 6 to 10). Through these momentum parameters, Algorithm 3 updates the model parameters for the current round for each of the aggregation startegies being considered (Algorithm 3: Line 11). The updated model parameters from the aggregation strategies are then grouped into a set (Algorithm 3: Line 12). Algorithm 3 computes the distances between the previous and current model parameters (Algorithm 3: Line 13), based on their Frobenius norm differences, and stores these distances in a separate set (Algorithm 3: Line 14). Finally, Algorithm 3 selects the most suitable aggregation strategy for the current round and returns the aggregated model parameters using it (Algorithm 3: Line 16). Algorithm 3 is directly integrated into Algorithm 1 by calling it (Algorithm 1: Line 22). 3.6 Conclusion This methodology section has provided a compact review of all the essential components of this study. The research framework visualizes the inputs and outputs at each stage. The method implementation outlines all the necessary steps to be taken. The datasets to be used have been introduced in detail. There has been a comprehensive exposition of the LICFL algorithm, the model parameter based cohorting algorithm, and the adaptive strategy selection algorithm. Therefore, adequate preparation has been made to proceed with the next phase of the study. Chapter 4 4.0 Work Plan and Timeline 4.1 Work Activities for Capstone 1 Work Activity 001 Introduction Description Writing the introduction chapter of the paper Work Products Draft of the introduction Revised introduction (after feedback) Scheduled Duration 2 weeks (10 working days) Risk Factors Overpromising Results Insufficient Context Weak Justification Unclear Objectives Decomposition Preparation (2 days) Composing (6 days) Refinement (2 days) Dependencies Research questions Research objectives Work Activity 002 Literature Review Description Writing the literature review chapter of the paper Work Products Draft of the literature review Revised literature review (after feedback) Scheduled Duration 4 weeks (20 working days) Risk Factors Lack of Focus or Relevance Bias in Selection Superficial Analysis Weak Connection to Research Goals Decomposition Preparation (5 days) Composing (10 days) Refinement (5 days) Dependencies Background of study Scope of study Work Activity 003 Methodology Description Writing the methodology chapter of the paper Work Products Draft of the methodology Revised methodology (after feedback) Scheduled Duration 2 weeks (10 working days) Risk Factors Insufficient Detail Incomplete Justification of Methodological Choices Over-Reliance on Secondary Sources Ignoring Limitations Decomposition Preparation (2 days) Composing (6 days) Refinement (2 days) Dependencies Problem statement 4.2 Gantt Chart for Capstone 1 4.3 Work Activities for Capstone 2 Work Activity 001 Feature Engineering Description Extract and select features from the datasets that represent key aspects of load prediction. Work Products Feature selection report Engineered datasets Scheduled Duration 1 week (5 working days) Risk Factors Insufficient data quality Irrelevant features Dependencies Preliminary problem statement and dataset analysis Work Activity 002 Data Preprocessing Description Preprocess the datasets by handling missing values, managing outliers, and applying normalization. Split the datasets into training and testing sets. Ensure the relevance of the testing set. Work Products Cleaned and preprocessed datasets Normalization/transformation scripts Training and testing datasets Scheduled Duration 1 week (5 working days) Risk Factors Incomplete handling of missing values Over-normalization that affects data variance Dependencies Preliminary understanding of dataset characteristics Defined criteria for dataset splits Work Activity 003 Model Construction Description Construct the LICFL algorithm, the model parameter based cohorting algorithm, and the adaptive strategy selection algorithm. Work Products LICFL algorithm implementation Cohorting algorithm implementation Adaptive strategy selection module Scheduled Duration 4 weeks (20 working days) Risk Factors Algorithmic complexity and inefficiency Poor integration of sub-algorithms Dependencies Availability of software development tools and libraries Work Activity 004 Model Training Description Configure model hyperparameters. Train the model using the training set. Record the training time. Work Products Trained model Hyperparameter tuning report Model performance log Training time metrics Scheduled Duration 2 weeks (10 working days) Risk Factors Poor hyperparameter optimization Long training times Inadequate model convergence Dependencies Access to training dataset Defined evaluation criteria for hyperparameters Work Activity 005 Model Inference Description Use the testing set to conduct inference with the trained model. Record the inference time. Work Products Inference results Inference time log Scheduled Duration 1 week (5 working days) Risk Factors Slow inference time Model not working as expected Dependencies Availability of trained model Testing dataset ready Work Activity 006 Comparative Experiment Description Select deep learning models as comparative models. Ensure the same datasets have been applied. Work Products Comparative model selection report Benchmark results Scheduled Duration 1 week (5 working days) Risk Factors Selection of inappropriate models for comparison Misalignment of experimental conditions Dependencies Availability of comparative models Clear definition of evaluation metrics Work Activity 007 Results Analysis Description Compare the performance of the models across the various metrics. Analyze their strengths and weaknesses. Work Products Performance comparison report Strengths and weaknesses analysis Visualizations of model performance Scheduled Duration 1 week (5 working days) Risk Factors Inaccurate analysis due to incomplete results Bias in performance evaluation Misinterpretation of metrics Dependencies Completion of comparative experiment Access to performance metrics Clear definition of evaluation criteria Work Activity 008 Conclusion Description Summarize the models' strengths and weaknesses that have been identified in the results analysis. Discuss the significance of the various metrics in load prediction. Provide recommendations for future improvements based on the findings. Work Products Final report Summary of strengths and weaknesses Recommendations for future work Conclusions drawn from metrics Scheduled Duration 1 week (5 working days) Risk Factors Incomplete or biased conclusions Lack of actionable recommendations Dependencies Results analysis report Clear understanding of metric importance 4.4 Gantt Chart for Capstone 2 5.0 References Abreha, H. G., Hayajneh, M., & Serhani, M. A. (2022). Federated learning in edge computing: A systematic survey. Sensors, 22(2), 450. https://doi.org/10.3390/s22020450 An, T., Ma, L., Wang, W., Yang, Y., Wang, J., & Chen, Y. (2023). Consideration of FedProx in privacy protection. Electronics, 12(20), 4364. https://doi.org/10.3390/electronics12204364 Ang, F., Chen, L., Zhao, N., Chen, Y., Wang, W., & Yu, F. R. (2020). Robust federated learning with noisy communication. IEEE Transactions on Communications, 68(6), 3452– 3464. https://doi.org/10.1109/tcomm.2020.2979149 Amarlingam, M., Wani, A., & NL, A. (2024). Lightweight industrial cohorted federated learning for heterogeneous assets. https://arxiv.org/abs/2407.17999 Arvanitidis, A. I., Bargiotas, D., Daskalopulu, A., Kontogiannis, D., Panapakidis, I. P., & Tsoukalas, L. H. (2022). Clustering informed MLP models for fast and accurate Short-Term Load forecasting. Energies, 15(4), 1295. https://doi.org/10.3390/en15041295 Bianchi, F. M., Maiorino, E., Kampffmeyer, M. C., Rizzi, A., & Jenssen, R. (2017). Recurrent neural networks for short-term load forecasting: An overview and comparative analysis. SpringerBriefs in Computer Science. https://doi.org/10.1007/978-3-319-70338-1 Biswal, B., Deb, S., Datta, S., Ustun, T. S., & Cali, U. (2024). Review on smart grid load forecasting for smart energy management using machine learning and deep learning techniques. Energy Reports, 12, 3654–3670. https://doi.org/10.1016/j.egyr.2024.09.056 Chai, Di & Wang, Leye & Chen, Kai & Yang, Qiang. (2020). FedEval: A benchmark system with a comprehensive evaluation model for federated learning. http://dx.doi.org/10.48550/arXiv.2011.09655 Dileep, G. (2020). A survey on smart grid technologies and applications. Renewable Energy, 146, 2589–2625. https://doi.org/10.1016/j.renene.2019.08.092 Fallah, S., Deo, R., Shojafar, M., Conti, M., & Shamshirband, S. (2018). Computational intelligence approaches for energy load forecasting in smart energy management grids: State of the art, future challenges, and research directions. Energies, 11(3), 596. https://doi.org/10.3390/en11030596 Farhadi, V., Mehmeti, F., He, T., La Porta, T., Khamfroush, H., Wang, S., & Chan, K. S. (2019). Service placement and request scheduling for data-intensive applications in edge clouds. IEEE INFOCOM 2019-IEEE Conference on Computer Communications, 1279–1287. https://doi.org/10.1109/infocom.2019.8737368 Fekri, M. N., Patel, H., Grolinger, K., & Sharma, V. (2020). Deep learning for load forecasting with smart meter data: Online Adaptive Recurrent Neural Network. Applied Energy, 282, 116177. https://doi.org/10.1016/j.apenergy.2020.116177 Fekri, M. N., Grolinger, K., & Mir, S. (2021). Distributed load forecasting using smart meter data: Federated learning with Recurrent Neural Networks. International Journal of Electrical Power & Energy Systems, 137, 107669. https://doi.org/10.1016/j.ijepes.2021.107669 Fekri, M. N., Grolinger, K., & Mir, S. (2023). Asynchronous adaptive federated learning for distributed load forecasting with smart meter data. International Journal of Electrical Power & Energy Systems, 153, 109285. https://doi.org/10.1016/j.ijepes.2023.109285 Felbab, V., Kiss, P., & Horváth, T. (2019). Optimization in federated learning. ITAT, 58– 65. http://ceur-ws.org/Vol-2473/paper13.pdf Fernández, J. D., Menci, S. P., Lee, C. M., Rieger, A., & Fridgen, G. (2022). Privacy-preserving federated learning for residential short-term load forecasting. Applied Energy, 326, 119915. https://doi.org/10.1016/j.apenergy.2022.119915 Gregor, K., Danihelka, I., Graves, A., Rezende, D., & Wierstra, D. (2015). DRAW: a recurrent neural network for image generation. International Conference on Machine Learning, 1462–1471. http://ece.duke.edu/~lcarin/Zhe10.2.2015.pdf Guo, M., Lv, R., Miao, Z., Fei, F., Fu, Z., Wu, E., Lan, L., & Wang, M. (2024). Load forecasting and operation optimization of ice-storage air conditioners based on improved deep-belief network. Processes, 12(3), 523. https://doi.org/10.3390/pr12030523 Haque, A., & Rahman, S. (2022). Short-term electrical load forecasting through heuristic configuration of regularized deep neural network. Applied Soft Computing, 122, 108877. https://doi.org/10.1016/j.asoc.2022.108877 Hasan, M. K., Habib, A. A., Shukur, Z., Ibrahim, F., Islam, S., & Razzaque, M. A. (2022). Review on cyber-physical and cyber-security system in smart grid: Standards, protocols, constraints, and recommendations. Journal of Network and Computer Applications, 209, 103540. https://doi.org/10.1016/j.jnca.2022.103540 Heaton, J. (2017). Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning. Genetic Programming and Evolvable Machines, 19(1-2), 305–307. https://doi.org/10.1007/s10710-017-9314-z Hong, T., Pinson, P., Fan, S., Zareipour, H., Troccoli, A., & Hyndman, R. J. (2016). Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond. International Journal of Forecasting, 32(3), 896– 913. https://doi.org/10.1016/j.ijforecast.2016.02.001 Hosein, S., & Hosein, P. (2017). Load forecasting using deep neural networks. IEEE PES Innovative Smart Grid Technologies Conference. https://doi.org/10.1109/isgt.2017.8085971 Hudson, N., Hossain, M. J., Hosseinzadeh, M., Khamfroush, H., Rahnamay-Naeini, M., & Ghani, N. (2021). A framework for edge intelligent smart distribution grids via federated learning. 2021 International Conference on Computer Communications and Networks (ICCCN). https://doi.org/10.1109/icccn52240.2021.9522360 Husnoo, M. A., Anwar, A., Hosseinzadeh, N., Islam, S. N., Mahmood, A. N., & Doss, R. (2023). A secure federated learning framework for residential Short-Term load Forecasting. IEEE Transactions on Smart Grid, 15(2), 2044–2055. https://doi.org/10.1109/tsg.2023.3292382 Janjua, J. I., Fatima, A., Athar, A., Khan, M. S., Rehman, Z. U., Javed, R., Abbas, T., & Zahra, I. (2023). Federated learning for prediction of energy consumption in edge-enabled smart cities with healthcare infrasturcture crossroads. Journal of Population Therapeutics & Clinical Pharmacology, 30(19). https://doi.org/10.53555/jptcp.v30i19.3649 Khan, A. R., Mahmood, A., Safdar, A., Khan, Z. A., & Khan, N. A. (2016). Load forecasting, dynamic pricing and DSM in smart grid: A review. Renewable and Sustainable Energy Reviews, 54, 1311–1322. https://doi.org/10.1016/j.rser.2015.10.117 Kim, Y. J., & Hong, C. S. (2019). Blockchain-based node-aware dynamic weighting methods for improving federated learning performance. 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS), 1–4. https://doi.org/10.23919/apnoms.2019.8893114 Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., & Zhang, Y. (2017). Short-term residential load forecasting based on LSTM recurrent neural network. IEEE Transactions on Smart Grid, 10(1), 841–851. https://doi.org/10.1109/tsg.2017.2753802 Kong, W., Dong, Z. Y., Hill, D. J., Luo, F., & Xu, Y. (2018). Short-term residential load forecasting based on resident behaviour learning. IEEE Transactions on Power Systems, 33(1), 1087–1088. https://doi.org/10.1109/tpwrs.2017.2688178 Kuster, C., Rezgui, Y., & Mourshed, M. (2017). Electrical load forecasting models: A critical systematic review. Sustainable Cities and Society, 35, 257–270. https://doi.org/10.1016/j.scs.2017.08.009 Li, C., Guo, Q., Shao, L., Li, J., & Wu, H. (2022). Research on short-term load forecasting based on optimized GRU neural network. Electronics, 11(22), 3834. https://doi.org/10.3390/electronics11223834 Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: challenges, methods, and future directions. IEEE

Signal Processing Magazine, 37(3), 50–60. https://doi.org/10.1109/msp.2020.2975749 Li, X., Huang, K., Yang, W., Wang, S., & Zhang, Z. (2020). On the convergence of FedAvg on non-IID data. International Conference on Learning Representations. https://www.openreview.net/pdf?id=HJxNAnVtDS Li, Z., Li, Y., Liu, Y., Wang, P., Lu, R., & Gooi, H. B. (2020). Deep learning based densely connected network for load forecasting. IEEE Transactions on Power Systems, 36(4), 2829–2840. https://doi.org/10.1109/tpwrs.2020.3048359 Majeed, U., & Hong, C. S. (2019). FLchain: Federated learning via MEC-enabled blockchain network. 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS), 1–4. https://doi.org/10.23919/apnoms.2019.8892848 McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. a. Y. (2017). Communication-efficient learning of deep networks from decentralized data. International Conference on Artificial Intelligence and Statistics, 1273– 1282. http://proceedings.mlr.press/v54/mcmahan17a/mcmahan17a.pdf Mughal, F. R., He, J., Das, B., Dharejo, F. A., Zhu, N., Khan, S. B., & Alzahrani, S. (2024). Adaptive federated learning for resource-constrained IoT devices through edge intelligence and multi-edge clustering. Scientific Reports, 14(1). https://doi.org/10.1038/s41598-024-78239-z 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. 2019 IEEE Symposium on Security and Privacy (SP). https://doi.org/10.1109/sp.2019.00065 Nguyen, D. C., Ding, M., Pham, Q., Pathirana, P. N., Le, L. B., Seneviratne, A., Li, J., Niyato, D., & Poor, H. V. (2021). Federated learning meets blockchain in edge computing: Opportunities and challenges. IEEE Internet of Things Journal, 8(16), 12806–12825. https://doi.org/10.1109/jiot.2021.3072611 Nishio, T., & Yonetani, R. (2019). Client selection for federated learning with heterogeneous resources in mobile edge. In Proceedings of the ICC 2019—2019 IEEE International Conference on Communications (ICC), 1–7. https://doi.org/10.48550/arXiv.1804.08333 Pandove, D., Goel, S., & Rani, R. (2018). Systematic review of clustering high-dimensional and large datasets. ACM Transactions on Knowledge Discovery From Data, 12(2), 1– 68. https://doi.org/10.1145/3132088 Peng, Y., Wang, Y., Lu, X., Li, H., Shi, D., Wang, Z., & Li, J. (2019). Short-term load forecasting at different aggregation levels with predictability analysis. 2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia), 3385–3390. https://doi.org/10.1109/isgt-asia.2019.8881343 Powell, J., McCafferty-Leroux, A., Hilal, W., & Gadsden, S. A. (2024). Smart grids: A comprehensive survey of challenges, industry applications, and future trends. Energy Reports, 11, 5760–5785. https://doi.org/10.1016/j.egyr.2024.05.051 Raza, M. A., Aman, M. M., Abro, A. G., Tunio, M. A., Khatri, K. L., & Shahid, M. (2022). Challenges and potentials of implementing a smart grid for Pakistan’s electric network. Energy Strategy Reviews, 43, 100941. https://doi.org/10.1016/j.esr.2022.100941 Real, A. C., Luz, G. P., Sousa, J., Brito, M., & Vieira, S. (2024). Optimization of a photovoltaic- battery system using deep reinforcement learning and load forecasting. Energy and AI, 16, 100347. https://doi.org/10.1016/j.egyai.2024.100347 Reisizadeh, A., Mokhtari, A., Hassani, H., Jadbabaie, A., & Pedarsani, R. (2019). FedPAQ: A communication-efficient federated learning method with periodic averaging and quantization. International Conference on Artificial Intelligence and Statistics, 2021– 2031. http://proceedings.mlr.press/v108/reisizadeh20a/reisizadeh20a.pdf Sendin, A., Peña, I., & Angueira, P. (2014). Strategies for power line communications smart metering network deployment. Energies, 7(4), 2377–2420. https://doi.org/10.3390/en7042377 Shalev-Shwartz, S., & Ben-David, S. (2015). Understanding machine learning: From foundations to algorithms. Cambridge University Press. Taik, A., & Cherkaoui, S. (2020). Electrical load forecasting using edge computing and federated learning. ICC 2020 - 2020 IEEE International Conference on Communications (ICC). https://doi.org/10.1109/icc40277.2020.9148937 Tang, X., Chen, H., Xiang, W., Yang, J., & Zou, M. (2021). Short-term load forecasting using channel and temporal attention based temporal convolutional network. Electric Power Systems Research, 205, 107761. https://doi.org/10.1016/j.epsr.2021.107761 Tayab, U. B., Zia, A., Yang, F., Lu, J., & Kashif, M. (2020). Short-term load forecasting for microgrid energy management system using hybrid HHO-FNN model with best-basis stationary wavelet packet transform. Energy, 203, 117857. https://doi.org/10.1016/j.energy.2020.117857 Tuballa, M. L., & Abundo, M. L. (2016). Smart grid technologies: A review of the current state and future directions. Renewable and Sustainable Energy Reviews, 59, 710–725. https://doi.org/10.1016/j.rser.2016.01.011 Wei, W., Liu, L., Wu, Y., Su, G., & Iyengar, A. (2021). Gradient-leakage resilient federated learning. 2021 IEEE 41st International Conference on Distributed Computing Systems (ICDCS). https://doi.org/10.1109/icdcs51616.2021.00081 Wen, X., Liao, J., Niu, Q., Shen, N., & Bao, Y. (2024). Deep learning-driven hybrid model for short-term load forecasting and smart grid information management. Scientific Reports, 14(1). https://doi.org/10.1038/s41598-024-63262-x Yazici, I., Beyca, O. F., & Delen, D. (2022). Deep-learning-based short-term electricity load forecasting: A real case application. Engineering Applications of Artificial Intelligence, 109, 104645. https://doi.org/10.1016/j.engappai.2021.104645 Ye, Y., Li, S., Liu, F., Tang, Y., & Hu, W. (2020). EdgeFed: Optimized federated learning based on edge computing. IEEE Access, 8, 209191– 209198. https://doi.org/10.1109/access.2020.3038287 Zheng, N. J., Xu, N. C., Zhang, N. Z., & Li, N. X. (2017). Electric load forecasting in smart grids using Long-Short-Term-Memory based Recurrent Neural Network. 2017 51st Annual Conference on Information Sciences and Systems (CISS), 1– 6. https://doi.org/10.1109/ciss.2017.7926112 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48