ESAIM 2024 – 2nd European Symposium on Artificial Intelligence in Manufacturing

Multifaceted Applications of Federated Learning: Beyond Neural Networks

Tatjana Legler[1,2,*], Vinit Hegiste[1], and Martin Ruskowski[1,2]

[1] Chair of machine tools and control systems, University of Kaiserslautern-Landau (RPTU), 67663 Kaiserslautern, Germany
[2] Innovative Factory Systems, German Research Center for Artificial Intelligence (DFKI), 67663 Kaiserslautern, Germany

[*] Corresponding author. Tel.: +49-631 205 5199;. E-mail: tatjana.legler@rptu.de

Abstract. The advent of Federated Learning (FL) has brought about a revolutionary change in the field of machine learning, enabling the decentralised training of models across a multitude of devices while simultaneously maintaining the confidentiality of the data. In contrast to conventional centralized methodologies, FL maintains the localisation of data, with only model updates being shared. This methodology enhances model generalisation and stability without compromising data sovereignty. A variety of machine learning techniques, including support vector machines (SVMs) and decision trees, can be effectively utilised within the context of FL frameworks. SVMs offer efficient solutions for classification tasks with minimal computational overhead, while decision trees provide interpretable models for both classification and regression. This paper explores the application of these methods in FL settings, highlighting their advantages and potential use cases in diverse industries, particularly in manufacturing. Furthermore, it discusses the integration of reinforcement learning with FL, emphasising its potential for enhancing intelligent and adaptable decentralised systems.

Keywords: Federated Learning · Machine Learning Algorithms · Privacy-preserving Techniques.

1 Introduction

Federated Learning (FL) represents a paradigm shift in machine learning by facilitating a decentralized learning approach [21]. Unlike traditional centralized learning models, where data is consolidated on a single server for training [6] [15], FL enables model training across multiple decentralized devices or nodes while keeping all the training data localized. One of the primary objectives of FL is to utilize a broad, heterogeneous database without the need to share sensitive or proprietary information [16]. This method ensures that only model updates, such as network weights, are shared between nodes, rather than raw data like sensor data or images [12]. Such an approach preserves data privacy and benefits all participants by combining the advantages of data sovereignty with the collaborative power of model training, thereby enhancing model generalizability and stability [30].

The introduction of Federated Averaging (FedAvg) by McMahan et al. [21] established a structured approach for applying FL through an iterative process. Initially, a central server initializes a global model and distributes it to all participating devices or nodes. Each node then independently trains the model on its own local dataset. After local training, nodes send their model updates, such as weights, back to the central server. The server aggregates these updates, typically by averaging, to update the global model. The updated global model is then sent back to the nodes. This process, known as a communication round, repeats until the model reaches the desired level of accuracy or meets other convergence criteria (see Figure 1). This iterative cycle leverages distributed data sources while maintaining the privacy and security of the data handled by each node.

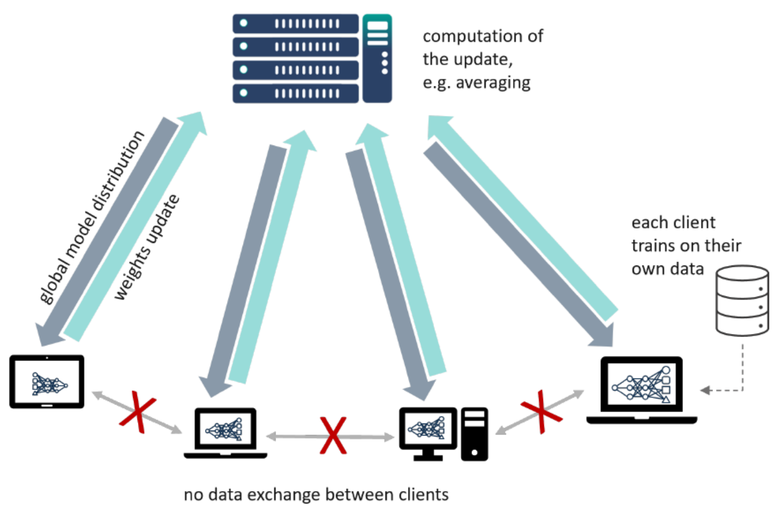

Fig. 1. This illustration depicts the federated learning process where a global model is distributed to multiple clients. Each client independently trains the model on their local data without exchanging data with other clients. The locally trained models' updates are sent back to a central server, which computes the aggregated update, typically through averaging. [18]

2 Machine Learning Methods in Federated Settings

While neural networks are known for their ability to handle complex and high-dimensional data, they are not always the most efficient or necessary approach for every problem. This is particularly true in FL settings, where the primary objective is to perform decentralized machine learning effectively and efficiently. Neural networks, especially deep learning models, are resource-intensive in terms of both computational power and data requirements [9]. Their high dimensionality increases communication overhead, potentially leading to inefficiencies and delays in the FL process [32].

Moreover, many real-world problems do not require the extensive modeling power of neural networks. Simpler approaches, such as Support Vector Machines (SVMs) or tree-based models, can often achieve comparable or sufficient accuracy with significantly less computational overhead and complexity [2]. SVMs are particularly well-suited for scenarios where the margin of separation and generalization to unseen data are more critical than capturing complex data patterns.

2.1 Support Vector Machines

SVM are a class of supervised learning algorithms used for classification and regression tasks. They are based on the concept of finding a hyperplane that best separates the data points of different classes in a high-dimensional space [26]. The main idea is to maximize the margin, which is the distance between the hyperplane and the closest data points from each class, known as support vectors. This margin maximization leads to better generalization and robustness of the classifier. SVMs can handle both linear and non-linear classification by using kernel functions, such as polynomial, radial basis, and sigmoid function, to map the input data into higher-dimensional spaces where it becomes linearly separable [17]. The theoretical foundations of SVMs are rooted in statistical learning theory, particularly in the concept of Structural Risk Minimization, which aims to find a balance between model complexity and fitting the training data to minimize the generalization error [29]. In a federated setting, the solution calculated by SVM must accommodate not only one dataset, but multiple subsets as accessible per client. This may result in a slightly different angle of the hyperplane, as illustrated in Figure 2.

The applications of SVM in manufacturing are numerous and diverse. The following are a few illustrative examples: In [1], SVMs are used to diagnose mechanical faults in motors by analyzing vibration signals collected from accelerometers. The SVMs classify the transformed vibration data to identify specific issues such as unbalance, misalignment, and mechanical looseness. In [28], SVMs predict anomalies within the manufacturing process. Upon detecting an anomaly, the system dynamically reconfigures itself to mitigate the issue, rerouting jobs to different machines or adjusting processing paths to maintain optimal operation and prevent overloading any single machine. In [19], SVMs are employed to

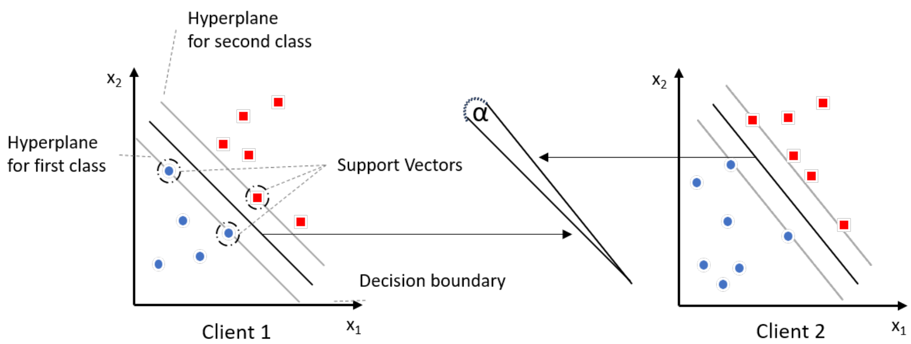

Fig. 2. Illustration of SVM Classifiers. The subfigures show a linear SVM classifier with a straight decision boundary and support vectors along the margins. The SVM aims to maximize the margin between the two classes, represented by red squares and blue circles.

predict manufacturing lead times. The model categorizes the total manufacturing time of products into different duration ranges, using production and work order data that undergo preprocessing and feature selection to enhance model performance.

A summary of the applications of SVM as referenced in the literature is presented herewith [25] [22]. By analyzing historical data, SVM models can identify the relationships between various input parameters and quality metrics, thereby enabling the prediction of product quality during production. By accurately categorizing products, SVMs assist in maintaining consistent quality standards and in the identification of batches that meet or fail to meet the required specifications. SVMs are also employed for the optimization of manufacturing processes, whereby the optimal settings for various machine parameters that result in the highest product quality are identified. Fault diagnosis, as previously mentioned, is also a key application area. By classifying the condition of machines as normal or faulty, SVMs enable the implementation of predictive maintenance strategies, which serve to reduce downtime and prevent catastrophic failures. This process entails training SVM models on historical data to ascertain the impact of varying parameter settings on the output quality. Manufacturers can subsequently utilize these models to simulate diverse scenarios and identify the optimal combination of parameters. To illustrate, in a machining process, SVMs can optimize cutting speed, feed rate, and tool geometry to minimize surface roughness and maximize material removal rate. This optimization not only enhances product quality but also improves process efficiency and reduces costs.

The aforementioned applications can be addressed and extended with FL, as the benefits are consistent across the board: It enhances SVM-based solutions by improving generalization by aggregating model updates from different units, creating a model that better predicts faults across different machines and environments. It enables scalable training across distributed locations, efficiently handling large datasets and allowing local models to quickly detect faults and take corrective action, reducing latency compared to centralized approaches. In addition, FL facilitates continuous learning, where local models are updated with new vibration data and these updates are periodically aggregated to refine the global model, ensuring that the system is always up-to-date with the latest patterns.

2.2 Decision Trees and Random Forrest

A Decision Tree (DT) is a widely used method in machine learning and data analysis for classification and regression tasks [5]. It operates by recursively splitting a dataset into subsets based on the most significant attribute, creating a tree-like structure where each internal node represents a test on an attribute, each branch corresponds to the test's outcome, and each leaf node signifies a final decision or prediction [24]. The selection of the best attribute to split on is typically determined by metrics such as Gini impurity, entropy, or information gain. They have also been employed to summarize associative classification rules, providing a more readable and compact classification mode [4]. They are valued for their simplicity, interpretability, and ability to handle both numerical and categorical data. However, they are prone to overfitting and can be unstable with small data variations. When employed in an ensemble, as is the case with random forest, the disadvantage of overfitting is mitigated [31].

In addition to similar areas of application as mentioned for SVM [23], DT can be used for decision support, e.g. for spare part configuration [3] as they are additionally better to interpret.[11] trained an incremental decision tree in a federated manner. For this they utilize 'Very Fast Decision Tree' (VFDT) as proposed in [7] and trained it in a vertical FL setting. [27] proposed a subtree-based horizontal FL method that accelerates model convergence and reduces communication costs while maintaining accuracy. Their FS-Boost approach learns one level of the tree at a time.

2.3 Reinforcement Learning

In distributed and decentralized systems, FL combined with reinforcement learning (FDRL) leverages the strengths of both methodologies to tackle complex, distributed learning tasks, such as in robotics. This powerful combination ensures data privacy while enabling seamless knowledge transfer between entities, thereby enhancing overall system intelligence and adaptability. FDRL frameworks, like FDRL, facilitate secure information sharing to build high-quality models while maintaining privacy protections [33]. Techniques such as reward shaping improve training efficiency and policy quality without compromising client confidentiality [14]. Furthermore, frameworks like Lifelong Federated Reinforcement Learning (LFRL) enhance robot navigation by fusing and transferring prior knowledge, allowing robots to quickly adapt to new environments [20]. This collaborative approach also supports multi-robot systems, enabling efficient task scheduling and improved learning performance across various applications

[13], [8]. In a modular factory setup, where each robot module is independent, as exemplified by the configuration depicted in [10], each module can be regarded as a FL client and train on their own objects when separated. However, upon connection, they share their updates and form a new global model. The potential for transferring learned skills across different entities using FDRL remains an open question, highlighting the need for further research in this area.

3 Conclusion and Outlook

This paper demonstrates the extensive potential applications of Federated Learning (FL) in manufacturing, with significant opportunities yet to be fully explored. Our exploration highlights numerous applications for Support Vector Machines (SVM) in areas such as defect detection, quality prediction, and parameter optimization. Additionally, Federated Distributed Reinforcement Learning shows considerable promise within robotics, particularly for path planning and grasp optimization. Our findings reveal that FL extends beyond neural networks, effectively incorporating a variety of machine learning methods. By combining different machine learning techniques with FL, robust, scalable, and privacy-preserving data analysis can be achieved. Each method offers specific strengths and alternatives that are computationally less demanding and more suited to particular types of data or tasks. The increasing awareness and concern over personal data usage, coupled with the enforcement of stringent data protection laws worldwide, underscore the necessity of privacy-preserving methods. By adhering to principles of data minimization and locality, FL aligns well with regulatory requirements and public sentiment, making it an indispensable tool in our data-driven world.

Future work will build on these findings, focusing on the combination of FL and SVM for planning optimization, and FL and Reinforcement Learning for robotics applications such as path planning and grasp optimization. This study underscores the versatility and potential of FL across a wide array of applications, paving the way for its broader adoption and integration into modern manufacturing environments. The continuous evolution of FL methodologies promises to further enhance data-driven decision-making processes while ensuring privacy and regulatory compliance.

Acknowledgment

This work was funded by the Carl Zeiss Stiftung, Germany under the Sustainable Embedded AI project (P2021-02-009).

References

1. Baccarini, L.M.R., Rocha e Silva, V.V., de Menezes, B.R., Caminhas, W.M.: Svm practical industrial application for mechanical faults diagnostic. Expert Systems with Applications **38**(6), 6980–6984 (2011)

2. Balabin, R.M., Lomakina, E.I.: Support vector machine regression - an alternative to neural networks for analytical chemistry? comparison of nonlinear methods on near infrared spectroscopy data. The Analyst **136**(8), 1703–1712 (2011)
3. Cantini, A., Peron, M., de Carlo, F., Sgarbossa, F.: A decision support system for configuring spare parts supply chains considering different manufacturing technologies. International Journal of Production Research **62**(8), 3023–3043 (2024)
4. Chen, Y.L., Hung, L.T.H.: Using decision trees to summarize associative classification rules. Expert Systems with Applications **36**(2), 2338–2351 (2009)
5. Costa, V.G., Pedreira, C.E.: Recent advances in decision trees: An updated survey. Artificial Intelligence Review **56**(5), 4765–4800 (2023)
6. Dean, J., Corrado, G., Monga, R., Chen, K., Devin, M., Mao, M., Ranzato, M.A., Senior, A., Tucker, P., Yang, K., Le, Q., Ng, A.: Large scale distributed deep networks. In: F. Pereira, C.J. Burges, L. Bottou, K.Q. Weinberger (eds.) Advances in Neural Information Processing Systems. vol. 25. Curran Associates, Inc (2012)
7. Domingos, P., Hulten, G.: Mining high-speed data streams. In: Ramakrishnan, R., Stolfo, S., Bayardo, R., Parsa, I. (eds.) Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 71–80. ACM, New York, NY, USA (2000)
8. Feng, W., Liu, H., Peng, X.: Federated reinforcement learning for sharing experiences between multiple workers. In: 2023 International Conference on Machine Learning and Cybernetics (ICMLC). pp. 440–445. IEEE (2023)
9. Frey, N.C., Li, B., McDonald, J., Zhao, D., Jones, M., Bestor, D., Tiwari, D., Gadepally, V., Samsi, S.: Benchmarking resource usage for efficient distributed deep learning. In: 2022 IEEE High Performance Extreme Computing Conference (HPEC). pp. 1–8. IEEE (2022)
10. Gafur, N., Kanagalingam, G., Wagner, A., Ruskowski, M.: Dynamic collision and deadlock avoidance for multiple robotic manipulators. IEEE Access **10**, 55766–55781 (2022)
11. Han, Z., Ge, C., Wu, B., Liu, Z.: Lightweight privacy-preserving federated incremental decision trees. IEEE Transactions on Services Computing **9**, 1–13 (2022)
12. Hegiste, V., Legler, T., Ruskowski, M.: Application of federated machine learning in manufacturing. In: 2022 International Conference on Industry 4.0 Technology (I4Tech). pp. 1–8. IEEE (2022)
13. Ho, T.M., Nguyen, K.K., Cheriet, M.: Federated deep reinforcement learning for task scheduling in heterogeneous autonomous robotic system. IEEE Transactions on Automation Science and Engineering **21**(1), 528–540 (2024)
14. Hu, Y., Hua, Y., Liu, W., Zhu, J.: Reward shaping based federated reinforcement learning. IEEE Access **9**, 67259–67267 (2021)
15. Iandola, F.N., Moskewicz, M.W., Ashraf, K., Keutzer, K.: Firecaffe: Near-linear acceleration of deep neural network training on compute clusters. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (2016)
16. Kairouz, P., McMahan, H.B., Avent, B., Bellet, A., Bennis, M., Nitin Bhagoji, A., et al: Advances and open problems in federated learning. FNT in Machine Learning (Foundations and Trends in Machine Learning) **14**(1–2), 1–210 (2021)
17. Kurani, A., Doshi, P., Vakharia, A., Shah, M.: A comprehensive comparative study of artificial neural network (ann) and support vector machines (svm) on stock forecasting. Annals of Data Science **10**(1), 183–208 (2023)
18. Legler, T., Hegiste, V., Ruskowski, M.: Mapping of newcomer clients in federated learning based on activation strength. 32nd International Conference Flexible Automation and Intelligent Manufacturing (2023)

19. Lim, Z.H., Yusof, U.K., Shamsudin, H.: Manufacturing lead time classification using support vector machine. In: Badioze Zaman, H., Smeaton, A.F., Shih, T.K., Velastin, S., Terutoshi, T., Mohamad Ali, N., Ahmad, M.N. (eds.) Advances in Visual Informatics, Lecture Notes in Computer Science, vol. 11870, pp. 268–278. Springer International Publishing, Cham (2019)
20. Liu, B., Wang, L., Liu, M.: Lifelong federated reinforcement learning: A learning architecture for navigation in cloud robotic systems. IEEE Robotics and Automation Letters **4**(4), 4555–4562 (2019)
21. McMahan, B., Moore, E., Ramage, D., Hampson, S., Arcas, B.A.y.: Communication-efficient learning of deep networks from decentralized data. In: Singh, A., Zhu, J. (eds.) Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. Proceedings of Machine Learning Research, vol. 54, pp. 1273–1282. PMLR, Fort Lauderdale, FL, USA (2017)
22. Paturi, U.M.R., Cheruku, S.: Application and performance of machine learning techniques in manufacturing sector from the past two decades: A review. Materials Today: Proceedings **38**, 2392–2401 (2021)
23. Plathottam, S.J., Rzonca, A., Lakhnori, R., Iloeje, C.O.: A review of artificial intelligence applications in manufacturing operations. Journal of Advanced Manufacturing and Processing **5**(3) (2023)
24. Rokach, L., Maimon, O.: Decision trees. In: Maimon, O., Rokach, L., Maimon, O.Z., Roaḥ, L. (eds.) Data Mining and Knowledge Discovery Handbook, pp. 165–192. SpringerLink Bücher, Springer US, Boston, MA (2005)
25. Rostami, H., Dantan, J.Y., Homri, L.: Review of data mining applications for quality assessment in manufacturing industry: support vector machines. International Journal of Metrology and Quality Engineering **6**(4), 401 (2015)
26. Saravanan, K., Prakash, R., Balakrishnan, C., Kumar, G.V.P., Siva Subramanian, R., Anita, M.: Support vector machines: Unveiling the power and versatility of svms in modern machine learning. In: 2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA). pp. 680–687. IEEE (2023)
27. Shimamura, K., Takamaeda-Yamazaki, S.: Fs-boost: Communication-efficient federated subtree-based gradient boosting decision trees. In: 2024 IEEE 21st Consumer Communications & Networking Conference (CCNC). pp. 839–842. IEEE (2024)
28. Shin, H.J., Cho, K.W., Oh, C.H.: Svm-based dynamic reconfiguration cps for manufacturing system in industry 4.0. Wireless Communications and Mobile Computing **2018**, 1–13 (2018)
29. Vapnik, Vladimir Naumovich and Vapnik, Vlamimir and others: Statistical learning theory. wiley New York (1998)
30. Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., Gao, Y.: A survey on federated learning. Knowledge-Based Systems **216**, 106775 (2021)
31. Zhang, M., Tao, F., Zuo, Y., Xiang, F., Wang, L., Nee, A.: Top ten intelligent algorithms towards smart manufacturing. Journal of Manufacturing Systems **71**, 158–171 (2023)
32. Zhang, W., Zhou, T., Lu, Q., Yuan, Y., Tolba, A., Said, W.: Fedsl: A communication-efficient federated learning with split layer aggregation. IEEE Internet of Things Journal **11**(9), 15587–15601 (2024)
33. Zhuo, H.H., Feng, W., Lin, Y., Xu, Q., Yang, Q.: Federated deep reinforcement learning