Journal scientific and applied research, vol. 29, 2025
International Journal

ISSN 1314-6289 (Print)
ISSN 2815-4622 (Online)

Original Contribution

MODEL ARCHITECTURE AND PERSPECTIVES OF FEDERATED LEARNING FOR SUSTAINABLE AGRICULTURE IN BULGARIA

Desislava N. Ivanova[1], Iliyana St. Ivanova[2]

[1]DEPARTMENT OF COMMUNICATION AND COMPUTER ENGINEERING AND SECURITY TECHNOLOGIES, FACULTY OF TECHNICAL SCIENCES, KONSTANTIN PRESLAVSKY UNIVERSITY OF SHUMEN, SHUMEN 9712,115, UNIVERSITETSKA STR., E-MAIL: d.n.ivanova@shu.bg
[2]DEPARTMENT OF COMMUNICATION AND COMPUTER ENGINEERING AND SECURITY TECHNOLOGIES, FACULTY OF TECHNICAL SCIENCES, KONSTANTIN PRESLAVSKY UNIVERSITY OF SHUMEN, SHUMEN 9712,115, UNIVERSITETSKA STR., E-MAIL: i.s.ivanova@shu.bg

ABSTRACT: *This paper presents a review of the application of Federated Learning (FL) in sustainable agriculture, with a specific focus on its potential integration within the Bulgarian agricultural sector. FL is a decentralized machine learning technology that enables multiple data sources to collaboratively train models without sharing raw data, thus preserving privacy and improving data security. The approach supports optimization of irrigation, fertilization, and pest control through locally adapted models and predictive analytics. The study analyzes the architecture, advantages, and limitations of FL, reviewing international case studies and exploring its relevance to Bulgarian conditions, such as regional climate variability and diverse soil types. Emphasis is placed on the potential role of FL in enhancing decision-making, reducing environmental impact, and supporting digital transformation in agriculture. Recommendations for implementation in Bulgarian farms and research institutions are also provided.*

KEY WORDS: *Federated learning, Sustainable agriculture, Data privacy, Model aggregation, Decentralized model training.*

1. Introduction

Sustainable agriculture is a strategic approach aimed at achieving a balance between economic efficiency, environmental sustainability and social equity. This approach aims not only to increase farm productivity but also to ensure the conservation of natural resources and improve the quality of rural life. Despite their importance, traditional methods of optimizing yields and managing agricultural processes remain limited by the lack of comprehensive, timely and relevant information.

Driven by the emergence of the Internet of Things (IoT) concept coupled with the massive use of sensors, drones and mobile devices, huge volumes of heterogeneous data - agronomic, climatic, soil and biological - are being generated. This data has the potential to modernize and fundamentally change agricultural practice through integration with artificial intelligence (AI) and machine learning (ML) based solutions. Through them, it is possible to achieve more accurate yield forecasting, better resource management, successful disease prevention and effective pest prevention.

However, centralized approaches to model training in AI, which require collecting all data in a single center, are accompanied by significant challenges. These include regulatory constraints related to the protection of personal and corporate data, privacy risks, and the need for stable and high-speed internet connectivity, which is a serious issue in rural areas.

In this regard, Federated Learning is establishing itself as a state-of-the-art solution in machine learning. It allows decentralized model learning, where data remains local to its owners (e.g. farms) and only model parameter updates are sent to a central server. FL not only preserves data confidentiality but also provides training capability under conditions of limited or unstable internet connectivity, making it particularly applicable in the agricultural sector.

At the international level, successful applications of Federated Learning (FL) in agriculture already exist. For example, in the Netherlands, a project utilizes FL to improve yield prediction and optimize irrigation strategies by leveraging sensor and weather data. In India, similar technologies are used to predict crop diseases, where farmers train local models without sharing sensitive information. In the U.S., under the PrecisionAg-FL initiative, FL supports monitoring of diseases in large agricultural areas through data from drones and IoT devices.

This paper examines the potential of using Federated Learning as a technological solution for the overall enhancement and improvement of sustainable agriculture in Bulgaria. It analyses the benefits and challenges of this approach, presenting applicable models, architectures and strategies for effective integration in the country's agricultural sector.

2. Theory and principles of federated learning

Blockchain Federated Learning is an approach where machine learning models are trained locally on multiple devices (clients) without the need for centralized data collection. The basic idea is to have each client train the model on its own data, then send the updated parameters to a central server for aggregation. This process can be described by the following formula [1]:

$$w_{t+1} = w_t + \eta \sum_{k=1}^{K} \left(\frac{n_k}{n} \Delta w_k\right) \quad (1)$$

Where:

- w_t - the global model at iteration t.
- η - the learning rate.
- K - the number of participating clients (farms).
- index of a client, к=1, …, K.
- n_k - the amount of data at client k.
- n - the total number of training examples across all clients.
- Δw_k - the local model change computed by client k.

This formula visualizes how the global model is updated by aggregating local updates from all clients. Each client contributes a change proportional to the number of its data relative to the total data in the system. This process is the basis of the Federated Averaging algorithm, which is a key mechanism in Federated Learning for creating a global model without compromising the confidentiality of local data.

Federated Learning has three main architectures-centralized, decentralized and hybrid [2] (Fig. 1).

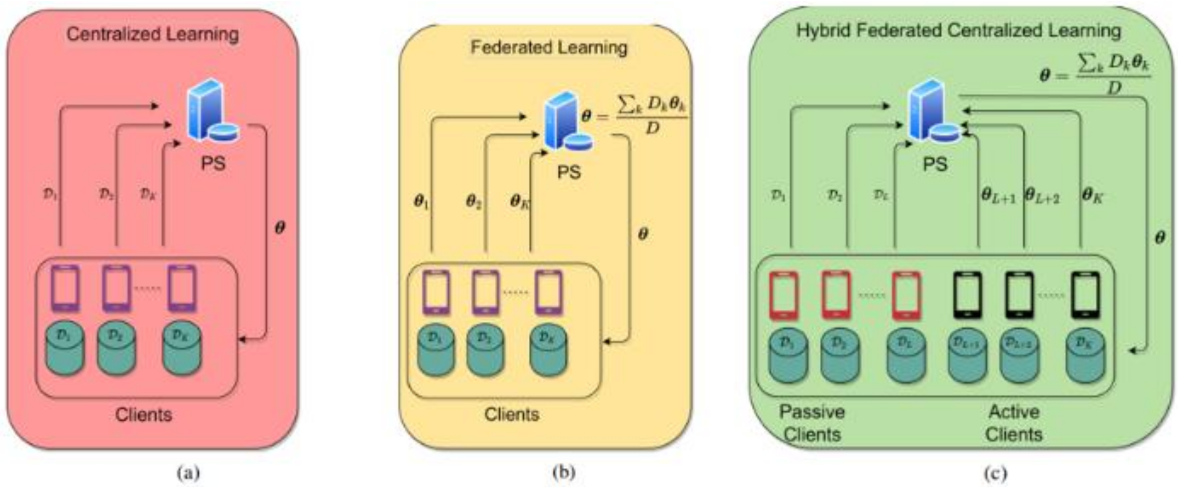

Fig. 1. CL, FL and HFCL frameworks. (a) In CL, all clients transmit their datasets to the PS. (b) In FL, the datasets are preserved at the clients while model parameters are sent to the PS. (c) In HFCL, the clients are designated as active and inactive depending on their computational capability to either perform CL or FL [2].

Centralized FL relies on a central server that coordinates learning by collecting parameters from participants and aggregating them into a global model [3]. Decentralized FL systems eliminate the need for a central server by directly sharing model parameters among clients in a peer-to-peer network [4]. The hybrid FL architecture combines centralized and federated learning - clients with sufficient computational resources participate in FL, while the rest send their data to a central server for traditional learning [5]. As described in the

Human-Centered AI Master's Programme (HCAIM) education portal [6], Federated Learning allows the modelling of a common model without the need for centralized collection of personal data, thus ensuring its privacy and security.

3. Federated Learning applications

The rapid digitization of agriculture has led to an unprecedented boom in data collection, necessitating the protection of privacy in innovative data analytics solutions. FL has emerged as a promising solution as it allows collaborative model learning across decentralized data sources without sharing raw data [7]. Table 1 presents the applications of Federated Learning in sustainable agriculture [7].

Table 1 FL applications in sustainable agriculture

Appendix	Description	Benefits for sustainable agriculture
Predicting the yield of different vegetable crops	Using FL to predict vegetable crop yields based on sensory and meteorological data	Improves harvest planning and optimization
Disease and pest detection in agricultural production	Analysis of images and other data by FL for early disease detection	Minimizing losses and reducing pesticide use
Optimization of irrigation	FL models are used for optimal irrigation management according to local conditions	Saving water and energy
Monitoring soil quality and nutrient availability	Collection and analysis of soil data for fertilizer recommendations	Reducing unnecessary fertilization and protecting the soil and the environment
Carbon footprint assessment	Using FL to assess and reduce greenhouse gas emissions	Supports sustainable agricultural practices

4. Using the prospects for the use of federated learning in sustainable agriculture in Bulgaria

In Bulgaria Federated Learning is still poorly applied in the agricultural sector, although the country has the prerequisites for its implementation. Bulgarian agriculture is diverse and covers important sectors such as viticulture, fruit growing and especially grain production, which forms a major part of the country's agrarian economy. The presence of a variety of agro-climatic conditions, different soil characteristics and regional practices makes Bulgaria a suitable terrain for FL, which precisely relies on decentralized learning on local specificities.

However, successful implementation of FL requires strategic action at the national level, such as:

- Creating a national alliance for digital farming involving universities, research institutes, farmers and technology companies.
- Incentives in the form of subsidies for farmers to implement IoT technologies, sensors and automation in agricultural production.
- Establishing contractual frameworks between actors (farmers, technology companies, institutions) that clearly define rights and responsibilities.
- Incorporating Federated Learning as a priority in research and education programs, especially in agricultural and engineering universities.
- Investment in network infrastructure and technologies such as 5G or satellite internet for rural areas.
- Using models and algorithms with fewer computational and communication requirements.
- Local edge devices that can perform a larger amount of computation autonomously (edge computing).

Investment in network infrastructure and technologies such as 5G or satellite internet for rural areas is essential. The security and management of real-time geospatial data are key aspects that correspond to the principles of Federated Learning in sustainable agriculture [8].

The authors of this paper propose a basic model for the application of Federated Learning in Bulgarian grain production.

The authors of this paper propose a basic model for the application of Federated Learning in Bulgarian grain production.

For the specific needs of this industry, they propose the use of a two-layer FL base model that combines different data sources and analysis algorithms. Table 2 presents the main components of the model for implementation in Bulgarian grain production.

Table 2 Main components of the Federated Learning Model

Component	Description
Local data and sensors	-Soil moisture, temperature, soil pH-measured by sensors to control conditions. -Meteorological data- rainfall, temperature. -Drone imaging- high resolution for crop condition diagnostics. -Satellite imagery - monitoring large areas through NDVI and other indices. -Historical data on yields and soil characteristics- basis

	for forecasting.
Local model of each farm	-CNN (Convolutional Neural Networks) - for drone and satellite imagery analysis. -LSTM (Long Short-Term Memory) -processing time series such as weather and soil data. -Decision Trees/Random Forests - for recommendations on seeding, irrigation and fertilizers.
Central Server	-Perform Federated Averaging to aggregate local models. - Updates the global model without raw data transfer. -Ensure security through Differential Privacy and Secure Aggregation.
Global Model	-Yield forecasts and risk assessment. -Disease warnings and need for actions. -Recommendations for irrigation, fertilization and agronomic practices. -Supports sustainable and efficient agriculture.

This model is a two-layer federated learning model, with a clear distinction between:

a. A local layer (first layer) includes the local models of each farm, which are trained on local sensor data, drone and satellite imagery, weather and historical data. Different models (CNN, LSTM, Decision Trees) are used here to process specific types of data collected in situ.
b. A central layer (second layer) comprising a central server that aggregates the local models via Federated Averaging (FedAvg) without collecting raw data. The central server maintains and updates the global model, which is used for forecasting and recommendations.

5. Interaction and connection of components in the system

Locally, each farm has a variety of sensors and data collection devices. These sensors measure critical parameters such as soil moisture, temperature, pH, and weather conditions such as rainfall and temperature. In addition, drone and satellite imagery provide visual information on crop status and development. All this data remains locally on the farm and is used to train local models - e.g. CNN for image analysis and LSTM for time series processing.

Once the local model is trained on the collected data, its parameters (not the data itself) are sent to the central server. This server collects updates from multiple farms in the area, allowing aggregation of models adapted to the specific climate and soil conditions of the region. Here, the Federated Averaging algorithm is applied to combine the local models into a more generalized model.

This process ensures that the raw data never leaves the farms while enabling knowledge sharing.

The local models are then sent to the central server, which merges them into a global model. The central server acts as a coordination hub - it updates and distributes the global model back to the local clients.

To ensure the confidentiality of the data and the security of the entire system, communication between all levels is accomplished by applying advanced security techniques - such as Differential Privacy, which adds noise to parameters, and Secure Aggregation, which allows servers to aggregate patterns without revealing individual updates. Ultimately, the global model provides farmers with yield forecasts, early disease warnings and recommendations for optimal agronomic practices. This supports sustainable agriculture in Bulgaria by improving efficiency and reducing negative environmental impacts. The architecture of the model used is presented in Fig. 2.

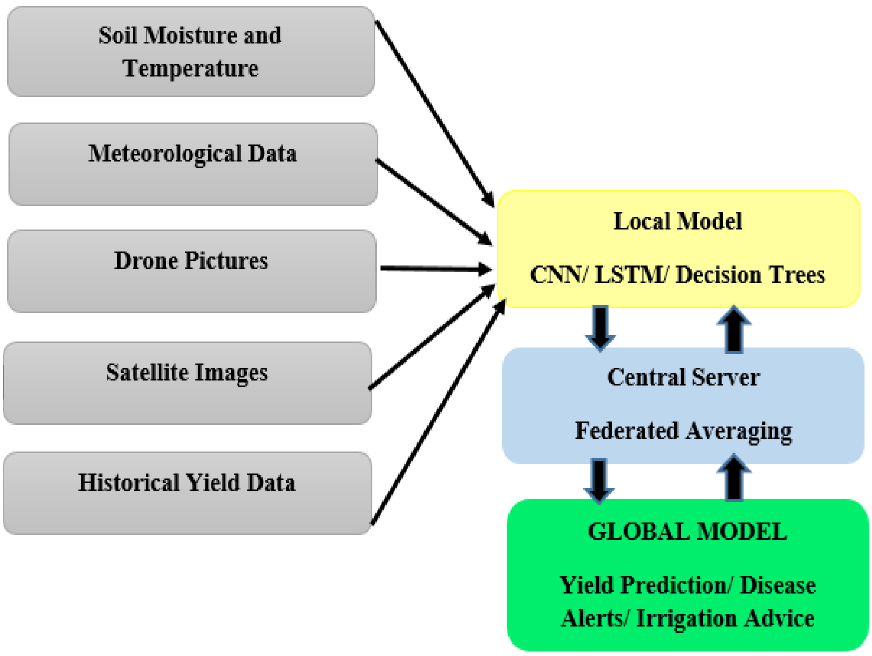

Fig. 2. Model for FL use in Bulgarian grain production

Similar models are already in use in the Netherlands and the USA. In the Netherlands, FL projects are being applied to predict potato and cereal yields, with integration of IoT sensors and drone imagery. In the US, within Precision Agriculture Initiatives, FL is used for early detection of crop stressors and yield optimization.

The use of FL models in Bulgarian agriculture is appropriate because they can be tailored to regional differences in different soils and climates. They allow the preservation of sensitive farm data and facilitate collaboration between small and large producers. Helps meet European requirements for sustainable agriculture and reduce carbon footprint.

As a member state of the European Union, Bulgaria follows key strategic documents related to digitalization, cybersecurity and data protection, such as Digital Decade 2030, GDPR, and the European Strategy for Data initiative, which creates a basis for the implementation of FL technology. It not only meets the requirements of GDPR but also supports the implementation of sustainable agriculture within the framework of European strategic objectives.

6. Discussion

The two-layer federated model, in which the first layer consists of locally trained models on farms, and the second layer consists of a central server that aggregates these models, offers several advantages in the context of Bulgarian agriculture, but also some challenges. One of its main advantages is the ability to protect sensitive agricultural data. Because raw data does not leave local devices, compliance with privacy requirements, including European regulations such as GDPR, is ensured.

Furthermore, the model allows adaptation to local agro-ecological conditions - each farm can train a model that considers specific characteristics such as soil type, microclimate, crops and practices. This is particularly important in a country like Bulgaria, where agronomic conditions vary even within the same region.

Another advantage is that the model can operate with limited connectivity, which is typical in some rural areas. Because the calculations are performed locally, permanent access to the Internet is not necessary.

The two-layer structure also offers good scalability - new actors (farms or even agri-districts) can be easily incorporated without changing the underlying architecture. An additional advantage is the use of different models according to data type - e.g. CNN for drone or satellite imagery, LSTM for time series from weather stations, and trees for tabular agronomic data.

However, the implementation of such a model in Bulgaria faces some difficulties. A problem is the lack of standardized technical infrastructure - there are currently no established platforms or examples of FL in agriculture in Bulgaria, which makes implementation expensive. In addition, data from different farms vary in volume and quality. This may lead to unequal representation of certain participants in the global model and result in skewed outcomes if importance-based and sample-size-based weighting mechanisms are not applied during aggregation. Another issue that may arise is related to the complexity of combining different model types in a central server - synchronization between heterogeneous architectures may require more advanced techniques or require custom models. A drawback is the lack of sufficient specialists in artificial intelligence, peripheral computing, and agricultural data analysis, which creates a barrier to the practical implementation of such an approach in most regions.

Despite these limitations, the two-layer federative model appears to be a good solution for introducing smart, sustainable agriculture in Bulgaria. With appropriate investments in technology, training and infrastructure, it can become the basis for more accurate, customized and sustainable agricultural management in the country.

To increase the efficiency, adaptability and applicability of the two-layer federative model in Bulgarian agriculture, it is advisable to focus its future development on the creation of a hierarchical (multilayer) federative architecture including an intermediate level of aggregation. In this extended structure, a regional hub (at agro-ecological region or district level) can be introduced that aggregates models from several nearby farms with similar climate and soil conditions before they are transmitted to the central server. Such architecture would improve the quality of the model, consider regional specificities, and reduce the burden on central infrastructure.

It is recommended to develop adaptive algorithms for local model weighting that consider not only the number of trained examples but also the quality and relevance of the data. This will allow a fairer and more efficient fusion of knowledge from diverse data sources.

In Bulgarian agriculture in particular, digital infrastructure is often inadequate. Therefore, it is important to invest in the development of local edge devices that can perform training even with limited connectivity. It is also necessary to build a national platform for federated agricultural training. This platform should allow farms, agronomists, and researchers to participate under standardized conditions and with access to shared resources.

These guidelines would contribute to the creation of a sustainable, secure and efficient agroecosystem based on smart technologies and decentralized knowledge, tailored to the realities and needs of the Bulgarian agricultural sector.

7. Conclusion

Federated Learning is an advanced and forward-looking technology that offers a pathway for the overall improvement of the agricultural sector - from traditional practices to smart, data-driven production management that is more sustainable, efficient and competitive. By combining local learning and global knowledge, FL enables the development of customized and confidential models capable of identifying the challenges of sustainable agriculture - from yield prediction and early disease detection to resource optimization and ecological footprint reduction.

Bulgaria has rich agro-climatic resources and scientific capacities that create preconditions for the implementation of FL in sectors such as grain production, viticulture and fruit growing. It is important to implement effective coordination between scientific institutions, farmers, technology companies and government authorities. This includes the establishment of national policies for digital farming, investment in IoT infrastructure and the training of personnel with competences in AI and agricultural sciences.

Acknowledgments

The author gratefully acknowledge the support provided by the project CoE UNITe BG16RFPR002-1.014-0004 funded by PRIDST, co-funded by the European Union. This study was supported by internal research project RD-08-109/05.02.2025 funded by Konstantin Preslavsky University of Shumen.

References:

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), 2017. Available: https://arxiv.org/abs/1602.05629.

[2] A. M. Elbir, S. Coleri, A. K. Papazafeiropoulos, P. Kourtessis, and S. Chatzinotas, "A Hybrid Architecture for Federated and Centralized Learning," IEEE Transactions on Cognitive Communications and Networking, vol. 8, no. 3, pp. 1529–1542, Sep. 2022.

[3] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in Proc. of AISTATS, pp. 1273–1282, 2017.

[4] A. Lalitha, O. C. Kilinc, T. Javidi, and F. Koushanfar, "Peer to Peer Federated Learning on Graphs," arXiv preprint arXiv:1901.11173, 2019.

[5] A. M. Elbir, S. Coleri, and K. V. Mishra, "Hybrid Federated and Centralized Learning," arXiv preprint arXiv:2011.06892, 2020.

[6] Human-Centered AI, "Theory of Federated Learning, Profiling and Personalization," https://humancentered-ai.eu/hcaim/lecture-theory-of-federated-learning-profiling-and-personalization-_BG.htm, accessed July 2025.

[7] R. Dembani, I. Karvelas, N. A. Akbar, S. Rizou, D. Tegolo, and S. Fountas, "Agricultural data privacy and federated learning: A review of challenges

and opportunities,” Computers and Electronics in Agriculture, vol. 203, p. 107509, 2023.
[8] Dimanova, D., Kuzmanov, Z., Real-time gis for monitoring and security, Journal Scientific and Applied Research, Vol. 27 No. 1 (2024), DOI: https://doi.org/10.46687/jsar.v27i1.408.