Federated Learning of Probabilistic Graphical Models

Pablo Torrijos[a,b,*]

[a]Instituto de Investigación en Informática de Albacete (I3A). Universidad de Castilla-La Mancha. Albacete, Spain.
[b]Departamento de Sistemas Informáticos. Universidad de Castilla-La Mancha. Albacete, Spain.

Abstract. Federated learning (FL) enables collaborative model training without data centralisation, yet most FL research targets deep neural networks. Interpretable models remain under-explored. This PhD investigates whether Bayesian networks (BNs), including Bayesian classifiers, can be learned federatively while maintaining three key properties: the privacy of local data, the tractability of the global model, and competitive predictive accuracy.

1 Introduction

Federated learning (FL) [8] enables multiple clients to collaboratively train a global model without sharing raw data. This paradigm aligns with modern privacy regulations, such as the General Data Protection Regulation (GDPR) [4], and mitigates practical barriers to data centralization. However, most FL research targets deep neural networks [20], offering little support for interpretability or probabilistic reasoning.

Bayesian networks (BNs) [6, 7] are probabilistic graphical models that encode conditional independencies via directed acyclic graphs. They offer explainability, uncertainty quantification, and robustness to missing data. These properties make them especially suitable for high-stakes decisions, aligning with the transparency and auditability requirements established by the European AI Act [3]. However, adapting BNs to the federated setting poses three key challenges:

- Q1 **Structure learning without data sharing.** How can clients independently learn BN structures and merge them into a global model without exchanging raw data?
- Q2 **Tractability and structural fidelity.** How can fusion operators on the server preserve relevant dependencies and maintain bounded model complexity without access to the data?
- Q3 **Formal privacy guarantees.** How can privacy be enforced during communication, whether of parameters or graphs, using mechanisms such as differential privacy [2] if necessary?

This PhD aims to extend the FL paradigm to probabilistic graphical models by developing algorithms for (i) distributed structure learning of BNs, (ii) federated classification with Bayesian models, and (iii) structure fusion under tractability and privacy constraints. The project addresses both fixed-structure classifiers, such as Naive Bayes [5] and Average n-Dependence Estimators (AnDE) [19]; fixed-structure classifiers, such as k-Dependence Bayesian (kDB) [10] and Tree Augmented Naive Bayes (TAN) [5]; and full BN learning via Greedy Equivalence Search (GES) [1]. In addition, it proposes novel structural fusion methods that aggregate graphs from different clients, limiting the unconstrained fusion [9] or creating consensus models that preserve key independencies.

The core hypothesis is that transparent, tractable, and privacy-preserving Bayesian models can be learned in a federated way without sacrificing predictive accuracy. Results to date support this claim, showing that federated BNs can match centralised baselines while satisfying formal privacy and scalability constraints.

* PhD student. Email: Pablo.Torrijos@uclm.es. Supervised by José M. Puerta and José A. Gámez. PhD started on 17th October 2022.

2 Work completed so far

The thesis has produced contributions in three areas: structure fusion, federated structure learning, and federated Bayesian classifiers. These contributions have been published or accepted in international venues and are supported by open-source implementations.

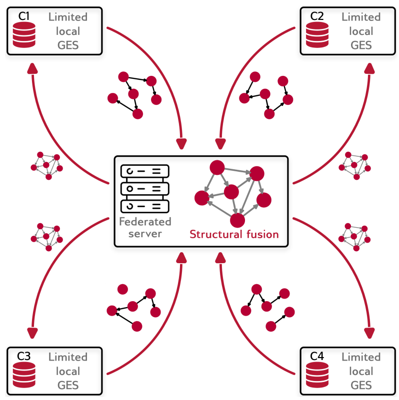

Figure 1. Diagram of FedGES [15], showing also the necessary component of structural fusion of Bayesian networks.

2.1 Tractable fusion of Bayesian network structures

We developed fusion strategies for aggregating heterogeneous Bayesian network (BN) structures while controlling model complexity. These contributions fall into three categories:

- **Fusion with direct treewidth constraint.** Genetic algorithms that optimise structural similarity while limiting the maximum structural complexity on the fused graph [13, 14].
- **Pre-fusion pruning.** A second family of genetic algorithms performs edge pruning before fusion finding for a consensus model after fusing the pruned BNs [16, 17].

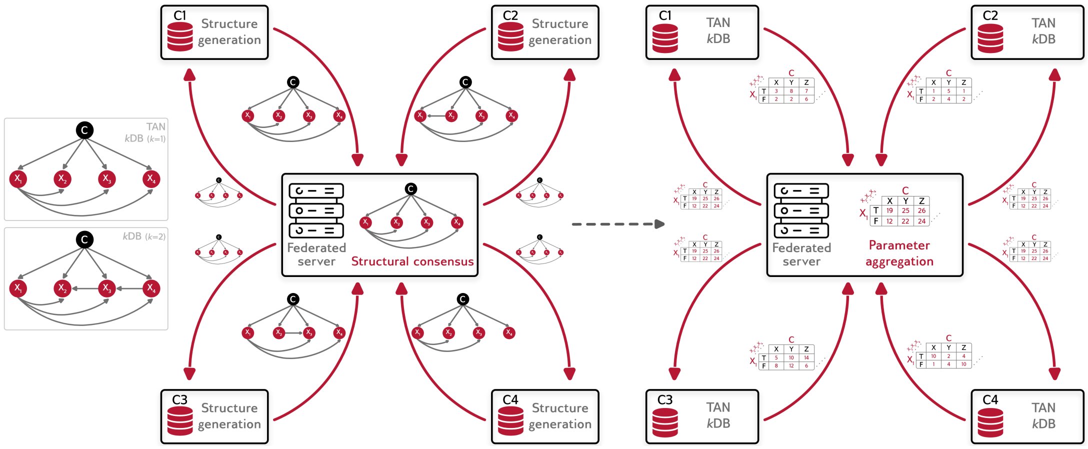

Figure 2. Diagram of federating Bayesian classifiers, specifically with variable structure, with the 2-step optimization.

- **Greedy consensus via minimum-cut.** A graph-based heuristic that integrates minimum-cut analysis into Backward Equivalence Search (BES) to select consensus structures without explicit thresholds [18].

These operators are compatible with both standalone fusion and iterative federated learning protocols, serving as core components in the construction of the global model.

2.2 *Federated learning of BN structures*

We proposed *FedGES* [15] (Figure 2), a protocol where clients learn BN structures using Greedy Equivalence Search (GES) and transmit only graphs to a central server. Multiple fusion operators can be employed at each round to construct a global structure, which is then utilized to guide the subsequent local search. An extended version of the method has been submitted to a journal.

2.3 *Federated Bayesian classifiers*

We adapted Naive Bayes and Average n-Dependence Estimators (AnDE) to the federated setting (Figure 2, right). In particular:

- *FedAnDE*, a framework for both generative and discriminative aggregation of AnDE classifiers (including Naive Bayes with $n = 0$) with support for differential privacy, has been submitted to a workshop and a journal.
- A discriminative version of federated Naive Bayes was previously published in [12].
- An earlier structural simplification of AnDE for high-dimensional data was presented in [11].

2.4 *Software and reproducibility*

All methods are released in the open-source library **BayesFL** library[1] (Java, MIT licence). This tool provides functionalities for federated structure learning of Bayesian networks and Bayesian classification, including multiple fusion and consensus algorithms. It features a modular architecture, designed to accommodate new algorithms and to support large-scale experimentation.

Regarding the datasets used in the experimental evaluation, we have published on **OpenML**[2] both the classification datasets not previously available in that repository, and the synthetic datasets generated from Bayesian networks in the *bnlearn* repository[3], sampled with 5000 instances per network. The latter are also available on **Zenodo**[4], with a total of **627** downloads.

3 Future work

The final phase of the thesis focuses on methodological extensions, integrating current components, and preparing journal versions. The main lines of work are:

F1 **Extension to variable-structure classifiers**. The fusion and consensus operators developed for structure-only models will be used as a preprocessing step to derive a common graph in scenarios with heterogeneous local structures (Figure 2). This will enable federated training of classifiers such as Tree-Augmented Naive Bayes (TAN) [5] and k-Dependence Bayesian (KDB) [10].

F2 **Journal extensions of published methods**. Three journal submissions are planned: (i) one combining the two families of genetic fusion methods (direct constraint and pre-fusion pruning); (ii) one focused on the greedy min-cut fusion method, including new greedy algorithm and expanded analysis; and (iii) an extended version of the MiniAnDE classifier, including its integration into the federated setting and evaluation on high-dimensional bioinformatics data. All submissions will include broader experiments and ablation studies.

F3 **Exploration of alternative local learners**. Although FedGES already provides convergence guarantees, the framework allows for other base structure learners. Future work includes testing alternative algorithms in the local step and analysing their impact on global model quality.

F4 **Application to real-world data**. The FedGES protocol, fusion methods, and federated Bayesian classifiers can be validated on large-scale, real-world healthcare datasets. The goal is to evaluate scalability and clinical interpretability in a realistic distributed scenario.

F5 **Thesis finalisation**. The dissertation will compile all published and submitted work and is expected to be completed and defended by December 2026, at the end of the four-year doctoral contract *FPU21/01074* funded by the Spanish Ministry of Science, Innovation and Universities.

[1] https://github.com/ptorrijos99/BayesFL
[2] https://www.openml.org/search?type=data&uploader_id=33148
[3] https://www.bnlearn.com/bnrepository
[4] https://doi.org/10.5281/zenodo.14917795

Acknowledgements

The following projects have funded this work: SBPLY/21/180225/000062 (Government of Castilla-La Mancha and ERDF funds); TED2021-131291B-I00, PID2022-139293NB-C32, and FPU21/01074 (MCIN/AEI/10.13039/501100011033 and ERDF Next Generation EU); 2025-GRIN-38476 (Universidad de Castilla-La Mancha and ERDF funds).

References

[1] D. M. Chickering. Optimal Structure Identification With Greedy Search. *Journal of Machine Learning Research*, 3:507–554, Jan 2002. ISSN ISSN 1533-7928.

[2] C. Dwork, F. McSherry, K. Nissim, and A. Smith. *Calibrating Noise to Sensitivity in Private Data Analysis*, page 265–284. Springer Berlin Heidelberg, 2006. doi: 10.1007/11681878_14.

[3] European Parliament and Council. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union, L1689, published 12 July 2024, 2024. URL https://eur-lex.europa.eu/eli/reg/2024/1689/oj.

[4] European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data (General Data Protection Regulation). Official Journal of the European Union, L119/1–88, published 4 May 2016, 2016. URL https://eur-lex.europa.eu/eli/reg/2016/679/oj.

[5] N. Friedman, D. Geiger, and M. Goldszmidt. Bayesian network classifiers. *Machine Learning*, 29:131–163, 1997. doi: 10.1023/a:1007465528199.

[6] F. V. Jensen and T. D. Nielsen. *Bayesian Networks and Decision Graphs*. Springer New York, 2nd edition, 2007. doi: 10.1007/978-0-387-68282-2.

[7] D. Koller and N. Friedman. *Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation and Machine Learning*. The MIT Press, 2009.

[8] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera. Communication-Efficient Learning of Deep Networks from Decentralized Data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pages 1273–1282. PMLR, Apr. 2017.

[9] J. M. Puerta, J. A. Aledo, J. A. Gámez, and J. D. Laborda. Efficient and accurate structural fusion of Bayesian networks. *Information Fusion*, 66:155–169, 2021. doi: 10.1016/j.inffus.2020.09.003.

[10] M. Sahami. Learning limited dependence bayesian classifiers. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, page 335–338. AAAI Press, 1996.

[11] P. Torrijos, J. A. Gámez, and J. M. Puerta. MiniAnDE: A Reduced AnDE Ensemble to Deal with Microarray Data. In *Engineering Applications of Neural Networks (EANN 2023)*, pages 131–143, Cham, Jan. 2023. Springer. doi: 10.1007/978-3-031-34204-2_12.

[12] P. Torrijos, J. C. Alfaro, J. A. Gámez, and J. M. Puerta. Federated Learning with Discriminative Naive Bayes Classifier. In *Intelligent Data Engineering and Automated Learning (IDEAL 2024)*, pages 328–339, Cham, Nov. 2024. Springer. doi: 10.1007/978-3-031-77738-7_27.

[13] P. Torrijos, J. A. Gámez, and J. M. Puerta. Structural fusion of bayesian networks with limited treewidth using genetic algorithms. In *2024 IEEE Congress on Evolutionary Computation (CEC)*, pages 1–8. IEEE, June 2024. doi: 10.1109/cec60901.2024.10611976.

[14] P. Torrijos, J. A. Gámez, and J. M. Puerta. Fusión Estructural de Redes Bayesianas con Treewidth Limitado utilizando Algoritmos Genéticos. In *Actas de la XX Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA 2024)*, pages 35–40. AEPIA, June 2024.

[15] P. Torrijos, J. A. Gámez, and J. M. Puerta. FedGES: A Federated Learning Approach for Bayesian Network Structure Learning. In *Discovery Science (DS 2024)*, pages 83–98, Cham, Jan. 2025. Springer. doi: 10.1007/978-3-031-78980-9_6.

[16] P. Torrijos, J. A. Gámez, J. M. Puerta, and J. A. Aledo. Algoritmos Genéticos para la Fusión Eficiente de Redes Bayesianas mediante Eliminación de Enlaces Previa a la Fusión. In *Actas del XVI Congreso Español de Metaheuristicas, Algoritmos Evolutivos y Bioinsipirado*, May 2025. URL https://openreview.net/forum?id=VDZcB9wl8m.

[17] P. Torrijos, J. M. Puerta, J. A. Gámez, and J. A. Aledo. Genetic Algorithms for Tractable Bayesian Network Fusion via Pre-Fusion Edge Pruning. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2025)*. ACM, July 2025. doi: 10.1145/3712256.3726333.

[18] P. Torrijos, J. M. Puerta, J. A. Gámez, and J. A. Aledo. Informed Greedy Algorithm for Scalable Bayesian Network Fusion via Minimum Cut Analysis, 2025.

[19] G. I. Webb, J. R. Boughton, and Z. Wang. Not so naive bayes: Aggregating one-dependence estimators. *Machine Learning*, 58(1):5–24, Jan. 2005. doi: 10.1007/s10994-005-4258-6.

[20] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao. A survey on federated learning. *Knowledge-Based Systems*, 216:106775, 2021. doi: 10.1016/j.knosys.2021.106775.