Combining Data Cleaning and Bias Mitigation in Federated Learning

Advisors

- Angela Bonifati (Lyon 1 Univ., angela.bonifati@univ-lyon1.fr)
- Sara Bouchenak (INSA Lyon, sara.bouchenak@insa-lyon.fr)

Duration

- 6 months

Context

Machine learning (ML) is applied in many areas to extract knowledge from data and guide the decision making process, with many applications such as search engines [7], recommendation systems [2] and disease diagnosis [6]. With the rapid growth of data, ML algorithms evolved from centralized to distributed solutions. And to address data privacy issues, Federated Learning (FL) has emerged to allow a set of participants to collectively resolve a machine learning problem without sharing their data.

Machine Learning plays a key role in decision making, therefore, it is crucial to ensure that the decisions made by a ML model do not reflect any discriminatory behaviour towards certain groups or populations. One of the factors that can lead to discriminatory decisions in ML is bias in the training data [12]. As defined in [13], bias is the inclination or prejudice of a decision made by a ML system which is for or against one person or group, in a way considered to be unfair. For instance, Amazon's recruiting tool was preferring male candidates over female candidates because the latter were under-represented in the training dataset [11].

Moreover, Federated Learning could exacerbate the problem of AI fairness and bias [8, 1]. Bias is a phenomenon that occurs when ML models produce unfair decisions due to the use of incomplete, faulty or prejudicial datasets and models. Bias may have serious consequences such as sexist segregation, illegal actions, or reduced revenues [3, 4, 9].

Federated Learning may have an impact on the problem of bias [8, 3], because of the decentralized nature of FL, where data distribution and size are particularly heterogeneous. Furthermore, data privacy constraints in FL do not allow the use of classical ML bias mitigation techniques [10, 5].

Assignment

The following research question is expected to be tackled during this internship : **How data preparation and cleaning affects the bias and fairness of Federated Learning algorithms ?** This raises the challenge of considering state-of-the-art approaches for data cleaning and profiling and studying their impact of a plethora of approaches targeting bias mitigation in FL.

The first objective is to identify relevant data quality problems on input data and available

open-source tools to fix or quantify the errors. Then, an exhaustive list of bias mitigation algorithms has to be built.
The second objective is to combine the two steps by considering the impact of data with better quality on the bias mitigation techniques, in a similar fashion to what is done with data quality and performance of classification models [15].

Tasks:

- Familiarize yourself with existing data cleaning techniques for raw data as well as bias mitigation in FL approaches
- Study the state of the art regarding the data repairing methods and data profiling (i.e. collecting statistics about the errors and inconsistencies in the data) as well as bias mitigation tools in FL
- Design and implement an experimental pipeline in which different data cleaning tools are combined with bias mitigation in FL tools in order to measure how quality impacts the bias mitigation and the FL model robustness and performance

This internship is funded by Action Transversale Liris between the Database team (Angela Bonifati) and the DRIM team (Sara Bouchenak).

Expected abilities:

- Very good programming skills (Python / C++/Java etc.)
- Very good communication skills
- Familiarity with data management/ML/distributed systems techniques

Opportunities: You will have the opportunity to work with top-class researchers in the above areas and to be possibly involved in writing a research paper.

References

1. Abay (A.), Chuba (E.), Zhou (Y.), Baracaldo (N.) et Ludwig (H.). – Addressing uniquefairness obstacles within federated learning. 2021.
2. Aher (S. B.) et Lobo (L.). – Combination of machine learning algorithms for recommen-dation of courses in e-learning system based on historical data.Knowledge-Based Systems,vol. 51, 2013.
3. Bellamy (R. K.), Dey (K.), Hind (M.), Hoffman (S. C.), Houde (S.), Kannan (K.), Lohia (P.),Martino (J.), Mehta (S.), Mojsilovic (A.) et al. – Ai fairness 360: An extensible toolkitfor detecting, understanding, and mitigating unwanted algorithmic bias.arXiv preprintarXiv:1810.01943, 2018.
4. Bogroff (A.) et Guegan (D.). – Artificial intelligence, data, ethics an holistic approach forrisks and regulation.University Ca'Foscari of Venice, Dept. of Economics Research Paper Series, no 19, 2019.
5. Calmon (F.), Wei (D.), Vinzamuri (B.), Natesan Ramamurthy (K.) et Varshney (K. R.). –Optimized pre-processing for discrimination prevention. – In Guyon (I.), Luxburg (U. V.),Bengio (S.), Wallach (H.), Fergus (R.), Vishwanathan (S.) et Garnett (R.) (édité par),Advances In Neural Information Processing Systems Volume 30. Curran Associates, Inc., 2017.
6. Kourou (K.), Exarchos (T. P.), Exarchos (K. P.), Karamouzis (M. V.) et Fotiadis (D. I.). –

Machine learning applications in cancer prognosis and prediction.Computational and structural biotechnology journal, vol. 13, 2015, pp. 8–17.
7. McCallumzy (A.), Nigamy (K.), Renniey (J.) et Seymorey (K.). – Building domain-specific search engines with machine learning techniques. – InProceedings of the AAAI Spring Sym-posium on Intelligent Agents in Cyberspace. Citeseer, pp. 28–39. Citeseer, 1999.
8. P. Kairouz et al. – Advances and open problems in federated learning.Foundations and Trends in Machine Learning, vol. 14, n1, 2021.
9. Wang (T.), Zhao (J.), Yatskar (M.), Chang (K.-W.) et Ordonez (V.). – Balanced datasets are not
enough: Estimating and mitigating gender bias in deep image representations. – InProceedings of the IEEE/CVF International Conference on Computer Vision, pp. 5310–5319, 2019.
10. Zemel (R.), Wu (Y.), Swersky (K.), Pitassi (T.) et Dwork (C.). – Learning fair representations.– In Dasgupta (S.) et McAllester (D.) (édité par),Proceedings of the 30th International Confer-ence on Machine Learning,Proceedings of Machine Learning Research, volume 28, pp. 325–333,Atlanta, Georgia, USA, 17–19 Jun 2013. PMLR.
11. Amazon scraps secret AI recruiting tool that showed biasa gainst women.–https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G
12. Calders (T.) et Zliobaite (I.). – Why unbiased computational processes can lead to discriminative decision procedures. vol. 3, 2013, pp. 43–57.
13. Ntoutsi (e. a.). – Bias in data-driven artificial intelligence systems—an introductory survey. WIREs Data Mining and Knowledge Discovery, vol. 10, 02 2020.
14. Wentai Wu, Ligang He, Weiwei Lin, Rui Mao, Chenlin Huang, Wei Song: FedProf: Optimizing Federated Learning with Dynamic Data Profiling. CoRR abs/2102.01733 (2021)
15. Lukas Budach, Moritz Feuerpfeil, Nina Ihde, Andrea Nathansen, Nele Sina Noack, Hendrik Patzlaff, Hazar Harmouch, Felix Naumann:
The Effects of Data Quality on ML-Model Performance. CoRR abs/2207.14529 (2022)

Selected Publications of Project Members:

- Laure Berti-Équille, Angela Bonifati, Tova Milo. Machine Learning to Data Management: A Round Trip. ICDE 2018: 1735-1738
- Remy Delanaux, Angela Bonifati, Marie-Christine Rousset, Romuald Thion. Query-Based Linked Data Anonymization. ISWC (1) 2018: 530-546
- Angela Bonifati, Radu Ciucanu, Slawek Staworko. Learning Join Queries from User Examples. ACM Trans. Database Syst. 40(4): 24:1-24:38 (2016)
- Angela Bonifati. Graph Queries. Generation, Evaluation and Learning (Invited Talk). EDBT/ICDT Workshops 2017
- B. Khalfoun, S. Ben Mokhtar, S. Bouchenak, V. Nitu. EDEN: Enforcing Location Privacy through Re-identification Risk Assessment: A Federated Learning Approach. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Volume 5, Issue 2, June 2021.
- M. Maouche, S. Ben Mokhtar, S. Bouchenak. HMC: Robust Privacy Protection of Mobility Data Against Multiple Re-Identification Attacks. ACM Journal on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2(3), September 2018.
- R. Talbi, S. Bouchenak, L. Y. Chen. Towards Dynamic End-to-End Privacy Preserving Data Classification. IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2018), Fast Abstract, Luxembourg, June 25-28, 2018.