The Role of Differential Privacy in GDPR Compliance

Position Paper

Rachel Cummings
Georgia Institute of Technology
School of Industrial and Systems Engineering
rachelc@gatech.edu

Deven Desai
Georgia Institute of Technology
Scheller College of Business
deven.desai@scheller.gatech.edu

ABSTRACT

The EU General Data Protection Regulation (GDPR) empowers individuals with the right to control erasure of their personal data held by firms. GDPR also allows firms to retain anonymized aggregate data and statistical results. Unfortunately, most recommender systems (and many other types of machine learning models) memoize individual data entries as they are trained, and thus are not sufficiently anonymized to be GDPR compliant. Differential privacy formally prevents against memoization and other types of overfitting, and additionally allows for accurate analysis in a wide variety of machine learning tasks. In this position paper, we advocate that differentially private learning should be the preferred method for GDPR-compliant recommender systems.

ACM Reference Format:
Rachel Cummings and Deven Desai. 2018. The Role of Differential Privacy in GDPR Compliance: Position Paper. In *Proceedings of Workshop on Responsible Recommendation (FATREC'18).* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 GDPR: PERSONAL DATA VERSUS STATISTICAL RESULTS

The EU General Data Protection Regulation (GDPR) specifies in Article 17 that individuals "shall have the right to obtain [...] the erasure of personal data concerning him or her without undue delay." However, GDPR makes different allowances for firms to retain aggregate data and statistical results. The notion of aggregate data is introduced in Recital 162 as part of the explanation of what it means to process personal data for statistical purposes. Recital 162 offers that "Statistical purposes mean any operation of collection and the processing of personal data necessary for statistical surveys or for the production of statistical results." A key point of Recital 162 is that the statistical results "may further be used for different purposes." This implies that the result of data processed for a statistical purpose is "aggregate data" as opposed to personal data, "and that this result or the personal data are not used in support of measures or decisions regarding any particular natural person." Given that machine learning is a statistical process and the outcomes are for statistical purposes, an open question is whether machine-learned models qualify as aggregate data.

As the former Chief Privacy Counsel of Microsoft has argued, under the GDPR aggregate data must also be anonymous. Specifically, such data must meet three criteria: (1) it must not be "directly linked to identifying data;" (2) there must not be a "known, systematic way to (re)identify the data; and (3) the data must not "relate to a specific person" [9]. This view connects to the language of Recital 26 which states, "The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable." Recital 26 concludes that the GDPR "does not therefore concern the processing of such anonymous information, including for statistical or research purposes." Thus for a machine learning model to qualify fully for exclusion from GDPR regulation, the model must meet the GDPR definition for anonymity.

In the alternative, given that anonymity is a high bar to meet and the exact definition for anonymity under GDPR is not settled, a model may not meet the anonymization standard and nonetheless be able to use pseudonymization under Article 4(5).[1] Pseudonymization is "processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person." Indeed, Article 89 which addresses safeguards and derogations relating to processing for statistical purposes, explicitly lists this technique as a measure that can be a safeguard "for the right and freedoms of data subjects."

2 MEMOIZATION OF PERSONAL DATA IN RECOMMENDER SYSTEMS

Many common machine learning algorithms *memoize* individual data entries during training, either intentionally by storing them in a cache to speed up processing, or inadvertently by imbedding personal data in the learned model. Carlini et al. [3] showed that deep learning algorithms for word prediction leaked Social Security Numbers and credit card numbers when trained on a corpus that included such data. Their results showed that "unintended memorization occurs early, is not due to over-fitting, and is a persistent issue across different types of models, hyperparameters, and training strategies" [3]. Earlier work [2] showed that collaborative filtering—the primary technique used in recommender systems—leaked information across users, where user i may learn personal data of user j through her personalized recommendations.

This memoization poses a problem for those who wish to have their model meet the GDPR's definition of statistical purpose and its safeguard standards of anonymization or pseudonymization. Since the underlying causes of inadvertent memoization are still poorly

FATREC'18, October 2018, Vancouver, Canada
2018. ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

[1] We thank Mike Hintze for his help in making this point.

understood—particularly in deep learning and neural networks [1]—it is difficult to ensure that a machine learned model is GDPR compliant, without using tools to formally prevent memoization.

3 THE PROMISE OF DIFFERENTIAL PRIVACY

In the last decade, *differential privacy* has emerged as the leading technique in computer science to allow for accurate data analysis with formal privacy guarantees. First defined by [6], differential privacy is a parameterized notion of database privacy that gives a mathematically rigorous worst-case bound on the maximum amount of information that can be learned about an individual's data from the output of a computation. It ensures the pair of outputs produced by two *neighboring databases* (which are the same except for one user's data) are nearly indistinguishable.

Definition 3.1 (Differential Privacy [6]). An algorithm $\mathcal{M} : \mathcal{D} \rightarrow \mathcal{R}$ is (ϵ, δ)-*differentially private* if for every pair of neighboring databases $X, X' \in \mathcal{D}$, and for every subset of possible outputs $\mathcal{S} \subseteq \mathcal{R}$,

$$\Pr[\mathcal{M}(X) \in \mathcal{S}] \leq \exp(\epsilon) \Pr[\mathcal{M}(X') \in \mathcal{S}] + \delta.$$

If $\delta = 0$, we say that $\mathcal{M}$ is ϵ-*differentially private.*

Differential privacy can be achieved for recommender systems by first privately learning the recommendation model, and then applying that model locally to each user's data for personalized recommendations. This ensures that for each user i, the set of personalized recommendations to all other users will be differentially private in i's data, but her own recommendation can be personalized to her data. See [7] for a textbook summary of differentially private algorithms that can be used to privately learn a recommendation model.

One critical benefit of differential privacy is that it provably prevents memoization. The requirement that private algorithms perform similarly on neighboring databases constrains the algorithm away from overfitting to individual entries in the database, and thus ensures that no single entry has been memoized. This guarantee also provides strong generalization guarantees for differentially private algorithms, which have also been observed in other machine learning applications [4, 5]. Empirically, Carlini et al. [3] implemented a differentially private version of their deep learning algorithm that previously leaked Social Security Numbers, and showed no such leakage occurred under the privately learned model. Further, their private word prediction model performed nearly as well as the non-private version.

4 FUTURE POLICY CHALLENGES

One major policy challenge in future implementations of differential privacy is determining appropriate values for the privacy parameters ϵ and δ. The primary privacy parameter is ϵ, where smaller ϵ corresponds to stronger privacy guarantees. Much of the theoretical literature advocates setting ϵ to be a small constant less than one, or to be diminishing in the size of the database (e.g., $O(1/\sqrt{n})$ for a database of size n) [10]. Practitioners prefer larger ϵ-values, because this weakening of privacy can yield improved accuracy of the data analysis. However, if ϵ is too large, the guarantees of differential privacy can cease to provide meaningful privacy guarantees. For example, recent work advocated that Apple's choice of parameters in their implementation of differential privacy provided insufficient privacy to users [11].

The second parameter δ is the maximum failure probability of the ϵ-differential privacy guarantee. Many commonly used differentially private algorithms set $\delta = 0$, with no possibility of failure. Using instead a cryptographically small positive value (i.e., $\delta = o(\exp(-n))$ for a database of size n) can lead to substantial improvements in accuracy. One method for GDPR-compliant data erasure is to encrypt personal data and delete the encryption key. If a cryptographically small failure probability is acceptable under GDPR for data erasures, then it may also be acceptable for anonymization of statistical results. The use of (ϵ, δ)-differential privacy may also be considered pseudonymization under Articles 4(5) and 89 of GDPR. Allowing a small $\delta > 0$ can significantly reduce the ϵ-value of an algorithm (via composition analysis [8]) which may yield overall privacy improvements.

REFERENCES

[1] Devansh Arpit, Stanisław Jastrzębski, Nicolas Ballas, David Krueger, Emmanuel Bengio, Maxinder S. Kanwal, Tegan Maharaj, Asja Fischer, Aaron Courville, Yoshua Bengio, and Simon Lacoste-Julien. 2017. A Closer Look at Memorization in Deep Networks. In *Proceedings of the 34th International Conference on Machine Learning (Proceedings of Machine Learning Research)*, Vol. 70. PMLR, 233–242.

[2] Joseph A. Calandrino, Ann Kilzer, Arvind Narayanan, Edward W. Felten, and Vitaly Shmatikov. 2011. "You Might Also Like:" Privacy Risks of Collaborative Filtering. In *2011 IEEE Symposium on Security and Privacy*. 231–246.

[3] Nicholas Carlini, Chang Liu, Jernej Kos, Ulfar Erlingsson, and Dawn Song. 2018. The Secret Sharer: Measuring Unintended Neural Network Memorization and Extracting Secrets. (2018). arXiv pre-print 1802.08232.

[4] Rachel Cummings, Katrina Ligett, Kobbi Nissim, Aaron Roth, and Zhiwei Steven Wu. 2016. Adaptive Learning with Robust Generalization Guarantees. In *29th Annual Conference on Learning Theory (COLT '16)*. 772–814.

[5] Cynthia Dwork, Vitaly Feldman, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Aaron Roth. 2015. The reusable holdout: Preserving validity in adaptive data analysis. *Science* 349, 6248 (2015), 636–638.

[6] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating noise to sensitivity in private data analysis. In *Proceedings of the 3rd Conference on Theory of Cryptography (TCC '06)*. 265–284.

[7] Cynthia Dwork and Aaron Roth. 2014. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science* 9, 3–4 (2014), 211–407.

[8] Cynthia Dwork, Guy N. Rothblum, and Salil Vadhan. 2010. Boosting and Differential Privacy. In *Proceedings of the IEEE 51st Annual Symposium on Foundations of Computer Science (FOCS '10)*. 51–60.

[9] Mike Hintze. 2018. Viewing the GDPR through a De-Identification Lens: A Tool for Compliance, Clarification, and Consistency. In *International Data Protection Law*, Vol. 8. 86–101.

[10] Justin Hsu, Marco Gaboardi, Andreas Haeberlen, Sanjeev Khanna, Arjun Narayan, Benjamin C. Pierce, and Aaron Roth. 2014. Differential Privacy: An Economic Method for Choosing Epsilon. In *IEEE 27th Computer Security Foundations Symposium*. 398–410.

[11] Jun Tang, Aleksandra Korolova, Xiaolong Bai, Xueqiang Wang, and Xiaofeng Wang. 2017. Privacy Loss in Apple's Implementation of Differential Privacy on MacOS 10.12. (2017). arXiv pre-print 1709.02753.