

Classification Protocols with Minimal Disclosure^{*}

Jinshuo Dong, Jason Hartline and Aravindan Vijayaraghavan

Northwestern University, Evanston, IL, USA

Abstract

We consider multi-party protocols for classification that are motivated by applications such as e-discovery in court proceedings. We identify a protocol that guarantees that the requesting party receives all responsive documents and the sending party discloses the minimal amount of non-responsive documents necessary to prove that all responsive documents have been received. This protocol can be embedded in a machine learning framework that enables automated labeling of points and the resulting multi-party protocol is equivalent to the standard one-party classification problem (if the one-party classification problem satisfies a natural independence-of-irrelevant-alternatives property). Our formal guarantees focus on the case where there is a linear classifier that correctly partitions the documents.

Keywords

classification, e-discovery, machine learning

preprint

ALTARS'23: 2nd Workshop on Augmented Intelligence for Technology-Assisted Reviews Systems: Evaluation Metrics and Protocols for eDiscovery and Systematic Review Systems. April 02, 2023, Dublin, Ireland

^{*} This work began during the IDEAL Special Quarter on Data Science and Law organized by Jason Hartline and Dan Linna. Many thanks to Dan Linna for legal context and feedback on the project. The work was supported in part by NSF award CCF-1934931. The recommended version of the paper is available at <https://arxiv.org/abs/2209.02690>

✉ jinshuo@northwestern.edu (J. Dong); hartline@northwestern.edu (J. Hartline); aravindv@northwestern.edu (A. Vijayaraghavan)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)