DISCUSSION PUBLISHERS
Research Center · Open Access Edition

Interpretable Machine Learning

Valerie Chen, Jeffrey Li, Joon Sik Kim, Gregory Plumb, Ameet Talwalkar

Discussion Research ID	DPR-2026-1145
Year	2021
Field	Computer Science
DOI	10.1145/3511299
Author ORCID	0009-0007-2783-0265
Licence	CC-BY
Original source	Queue
Permanent link	https://discussionpublishers.com/research/bba7ef85-a5cd-4f7c-96ea-c083f4a9aba3
Publisher	Discussion Publishers discussionpublishers.com

Republished by Discussion Publishers under the original CC-BY licence. Copyright remains with the original authors; please cite the original work.

Cite as: Valerie Chen, Jeffrey Li, Joon Sik Kim, Gregory Plumb, Ameet (2021). Interpretable Machine Learning. Discussion Publisher

Abstract

The emergence of machine learning as a society-changing technology in the past decade has triggered concerns about people's inability to understand the reasoning of increasingly complex models. The field of IML (interpretable machine learning) grew out of these concerns, with the goal of empowering various stakeholders to tackle use cases, such as building trust in models, performing model debugging, and generally informing real human decision-making.

Keywords: model interpretability, model explainability, model debugging, human decision-making

Full text

The complete text of this work is hosted by its original publisher.
Read it via the Discussion Publishers page:
https://discussionpublishers.com/research/bba7ef85-a5cd-4f7c-96ea-c083f4a9aba3

Original record: https://doi.org/10.1145/3511299
DOI: https://doi.org/10.1145/3511299