From: Paolo Giudici
To: ai-inquiries
Cc: Paolo Giudici
Subject: NIST Calls for Information to Support Safe, Secure and Trustworthy Development and Use of Artificial Intelligence
Date: Friday, December 22, 2023 12:22:36 PM

Dear Madam/Sirs,
with reference to the recent NIST Call for Information to Support Safe, Secure and Trustworthy Development and Use of Artificial Intelligence, please find below my comment.

I believe that, for an effective implementation of a Safe , secure and Trustworthy AI, it is necessary to operationalise AI risk management in a set of integrated metrics that can be employed to determine tolerance levels, in function of the risk appetite of the stakeholder involved.

This proposal is in strong analogy with what in use for financial risk management, where metrics based on the Value at Risk (VaR) are being used to measure market, credit, operational and other risks and, consequently, define tolerance levels in function of the risk appetite of the involved financial institution.

To actually operationalise AI risk management, I recently proposed, with other academics, an approach called S.A.F.E. AI which is based on a set of metrics based on the Lorenz concentration curve of the machine learning output, referred, respectively, to the measurement of Security (Sustainability), Accuracy, Fairness and Explainability of the AI output. The approach leads to the calculation of a normalised score for each of the four dimensions, each of which can be used to set a tolerance level. The scores can also be integrated in a unified measure of Safe and Trustworthy AI.

More technical details on the proposal can be found in two recently published papers, available open access:

https://doi.org/10.1016/j.frl.2023.104088

https://doi.org/10.1016/j.eswa.2023.121220

I am available to discuss further developments/extensions of the methodology, and its application to different industry (so far we mainly focused on the financial sector)

Looking forward to your acknowledgement of my comment above, and on its posting, I wish you all the best for the coming Christmas time.

best regards

Paolo Stefano Giudici
Professor of Statistics, Department of Economics and Management

Personal page: https://sites.google.com/a/unipv.it/giudici/
Stats Lab page: https://sites.google.com/unipv.it/statslab-pavia/home?authuser=0

---------- Forwarded message ---------
Da: **National Institute of Standards and Technology (NIST)** <subscriptions@service.govdelivery.com>
Date: mar 19 dic 2023 alle ore 23:21
Subject: NIST Calls for Information to Support Safe, Secure and Trustworthy Development and Use of Artificial Intelligence
To: <giudici@unipv.it>

View as a Web Page

News

NIST Calls for Information to Support Safe, Secure and Trustworthy Development and Use of Artificial Intelligence

The U.S. Department of Commerce’s National Institute of Standards and Technology (NIST) has issued a Request for Information (RFI) that will assist in the implementations of its responsibilities under the recent Executive Order on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (AI). The order directs NIST to develop guidelines for evaluation, red-teaming and more; facilitate development of consensus-based standards; and provide testing environments for the evaluation of AI systems. These guidelines and infrastructure will be a resource to help the AI community in the safe and trustworthy development and responsible use of AI.

“President Biden has been clear — AI is the defining technology of our generation, and we have an obligation to harness the power of AI for good while protecting people from its risks. As part of the president’s Executive Order, the Department of Commerce is soliciting feedback across industry, academia, civil society and more so we can develop industry standards around AI safety, security, and trust that will enable America to continue leading the world in the responsible development and use of this rapidly evolving technology," said U.S. Secretary of Commerce Gina Raimondo.

Read More

IN CASE YOU MISSED IT

NIST Offers Draft Guidance on Evaluating a Privacy Protection Technique for the AI Era

Dec. 11, 2023

The agency has made progress on one of its tasks delineated in the recent Executive Order on AI.

Read More

Connect with us

Received this email from a friend? Subscribe here.

NIST