Incorporating FAT,privacy aware AI into decision-making frameworks

26th January 2022

ASHISH.MEHTA,CSA BLOCKCHAIN WG

Stunning attacks on crypto ecosystem in January…

- Opensea (NFT marketplace)-$13 Billion valuation ;over a million dollars stolen in January via security flaws

  https://twitter.com/crypto_bitlord7/status/1485538722229944322

  https://www.zerohedge.com/markets/opensea-bug-allows-hackers-steal-more-1-million-nfts

- Multichain(cross-blockchain router protocol- $8 Billion valuation);underlying base for WETH, PERI, OMT, WBNB, MATIC& AVAX swaps;over $3 million stolen

  https://www.zdnet.com/article/multichain-token-hack-losses-reach-3-million-report/

- Cheap malware(redline,crypobot) lowers the barriers for entry for attacking cryptocurrency wallets-Cryptocurrency compliance teams need to ensure that malware isn't hitting them

  https://www.zdnet.com/article/cheap-malware-is-behind-a-rise-in-attacks-on-cryptocurrency-wallets/

Special Thanks to the three authors…

Dmitry Zhdanov , Sudip Bhattacharjee & Mikhail A. Bragin from University of Connecticut for sharing this paper with me & permitting me to share with this Working Group.

https://doi.org/10.1016/j.dss.2021.113715

Background to the research paper

- Researchers from Georgia State & University of Connecticut have built a formal approach to tracking Fairness,Accountability & Transparency across AI systems
- Added a privacy constrained toolset & built a model to balance F-A-T
- They have been able to effectively prove(via modeling) that AI &FAT can co-exist.
- Use-cases in insurance, health diagnostics, government funds allocation and other business settings.
- policy implications(from EU & Singapore) as well, by making AI and AI-based decisions more ethical, less controversial, and hence, trustworthy

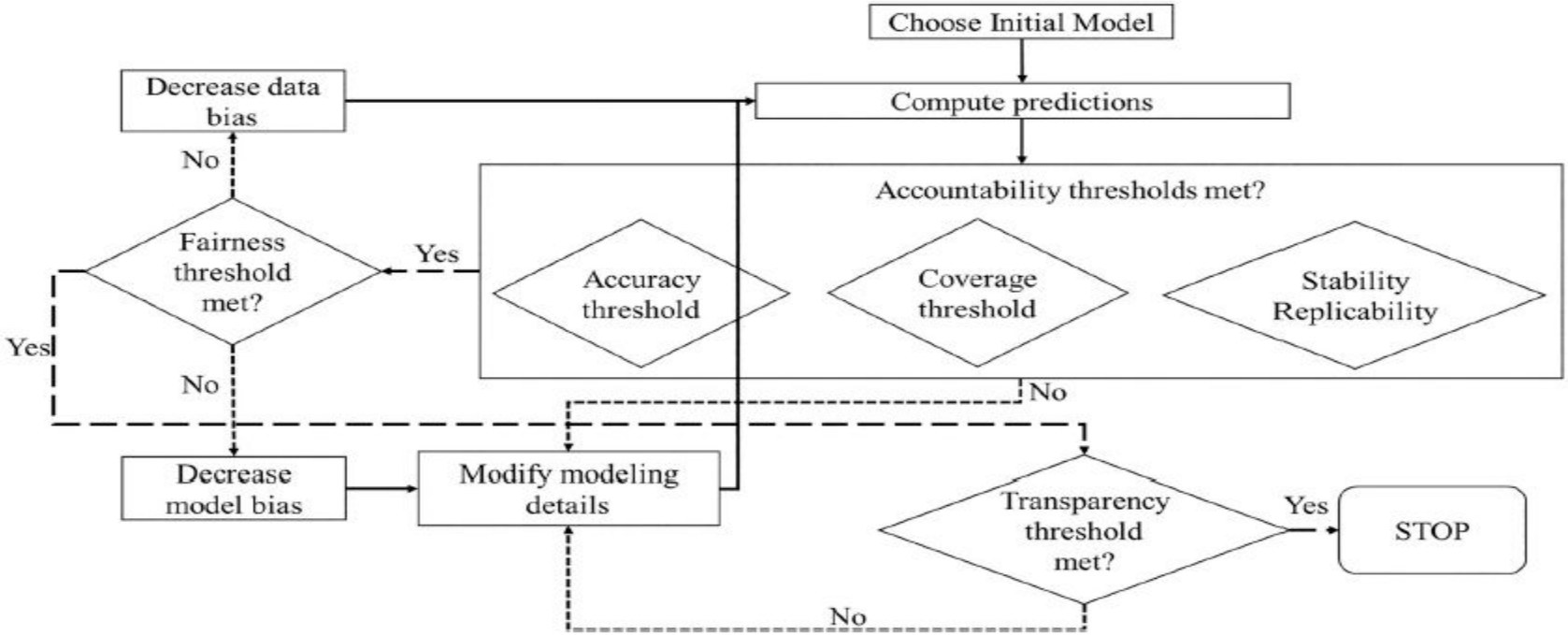
As AI gains prominence in Business..new issues arise

- AI need to be trusted & accepted by the people they impact as well as to follow regulations.
- traditionally focused on system performance and accuracy which may be highly precise but difficult to explain and interpret for decision making.
- Further, such systems may be biased either via data labels or systemic social biases), and possibly unethical-Question of AI ethics like recently in the case of major insurance judgements or how Amazon recruiting policies were against women.
- Furthermore how do you take into account privacy rules like GDPR or CCPA.

Six Principles of responsible AI

- AI need to be trusted & accepted by the people they impact as well as to follow regulations.
- Fairness, Accountability,Transparency, Privacy, Security, and Ethics.
- Authors show its possible to work with a Privacy constrained dataset (maintaining anonymity of key identifiers/low attribute data) while still delivering FAT in AI/ML.
- From the end-user perspective individual privacy concerns may limit trust and acceptance of technology such as AI-based systems.
- From the designer perspective, raise the issue of privacy in the presence of data fusion (where data comes from multiple sources and/or processed by multiple entities).

Computational model used for FAT



Instantization model used for affinity modelling

Instantization model used for affinity modelling

Performance of their model in a real-world scenario

Evaluation of performance in relation to FAT framework.

Accountability			Fairness	Transparency
Accuracy (RMSE vs benchmark)	Coverage	Stability	# Partitions	Model for prediction
0.9834 (−3.25%)	88.21%	Yes	4	Yes
0.9719 (−2.04%)	86.03%	Yes	6	Yes
0.9777 (−2.65%)	81.42%	Yes	11	Yes

Conclusions from their paper

- Accountability-Since they do not observe any dominant relationship between partitions, they conclude that our methodological framework is balanced and suitable for a wide variety of partitions and intrinsic product characteristics.
- Transparency-A low number of variables involved in the models facilitates the goal of transparency and explainability in decision making. Further, our ability to understand the impact of each predictor variable on accuracy and coverage.
- Fairness-They conclude that partitioning of large datasets into finer segments helps improve prediction accuracy, which indirectly points to increasing fairness across diverse segments of items.

Q&A