

Résumé de la bibliographie

7 novembre 2025

- M. Hekmatnejad et al.. (2024). *Formalizing and evaluating requirements of perception systems for automated vehicles using spatio-temporal perception logic*
Monitoring des algos de vision dans les véhicules autonomes. Section V : beaucoup d'exemples de propriétés à monitorer. (2025-09-17)
- D. Kang et al.. (2020). *Model Assertions for Monitoring and Improving ML Models*
P. Mallozzi, P. Pelliccione, P. P. A. C. Menghi. (2018). *Keeping intelligence under control*
Monitoring d'invariants sur le fonctionnement d'un modèle ML. Propose l'idée mais pas une solution. (2025-09-17)
- N. Aréchiga et al.. (2019). *Better AI through Logical Scaffolding*
Introduction de l'idée de "logical scaffolding"; encore une fois, davantage une ébauche qu'une idée aboutie. (2025-09-17)
- S. Bensalem et al.. (2024). *Bridging formal methods and machine learning with model checking and global optimisation*
Approche très mathématique : cherche à faire du model checking sur un algo d'IA. Par contre la section 5 montre plusieurs exemples intéressants de propriétés qu'on pourrait *monitorer*.
- J. López, M. Labonne, M. L. A. C. Poletti. (2021). *Toward Formal Data Set Verification for Building Effective Machine Learning Models*
— Utilise MSFOL pour exprimer des contraintes qui peuvent être satisfaites ou non par un dataset. - Partie intéressante : §3.2. Donne des exemples de propriétés : - Il doit y avoir au moins n instances - Le nombre d'instances de chaque classe est équilibré - Les données sont min-max normalisées (toutes entre x et y , et ces deux bornes sont atteintes) - Les instances échantillonnent adéquatement l'espace (chaque point est à une distance $< d$ de ses voisins) Curieusement, utilisent Z3 pour vérifier ces contraintes sur un dataset. Autrement dit, plutôt que de vérifier $\varphi(M)$, font plutôt $M \wedge \varphi$ –ce qui est "overkill".
- J. Guérin et al.. (2023). *Out-of-Distribution Detection Is Not All You Need*
Faire du snowballing sur cette ref, qui cite sûrement d'autres travaux sur le monitoring du "out-of-distribution".
- C. Qian et al.. (2022). *Distinguish Sense from Nonsense : Out-of-Scope Detection for Virtual Assistants*
P. Mallozzi, P. Pelliccione, P. P. A. C. Menghi. (2018). *Keeping intelligence under control*