

## Intelligence Artificielle (IA), algorithmes et machine learning

### Pense-bête de la CER-VD (v1.0, 26.10.2023)

1. Justification du recours à l'IA :
  - a. Pour quelles raisons le recours à l'IA (apprentissage statistique) est-il justifié par rapport aux objectifs du projet de recherche ? Est-ce que les conditions d'utilisation de l'IA sont réunies ?
  - b. Existe-t-il d'autres approches méthodologiques pour atteindre les objectifs fixés avec les ressources à disposition ?
2. Justification du modèle d'IA envisagé :
  - a. Comment le choix du modèle est-il justifié ?
  - b. Comment sont/seront constituées les vérités terrain (ground truth) ?
  - c. Quels sont les critères d'évaluation des performances utilisés ?
  - d. Quels sont les implications des différents types d'erreurs (erreurs du classifieur)?
  - e. Quelles mesures sont mises en place pour anticiper et prévenir les discriminations, biais et erreurs potentielles et les identifier <sup>1 2</sup>?
  - f. Quelles mesures sont mises en place pour vérifier la fiabilité des résultats (contrôle qualité)?
3. Confidentialité et sécurité des données :
  - a. D'où proviennent les données d'entraînement ? Avec quel consentement ?
  - b. Qui a accès aux données brutes et transformées (modèles appris, statistique)?
  - c. Comment les attaques par inférence d'appartenance (membership inference attack) sont-elles prévenues lors du partage des modèles appris ?
4. Responsabilités :
  - a. Comment est envisagée la transmission aux patient.e.s /participant.e.s et autres personnes concernées des résultats issus des processus d'IA (transparence et interprétabilité)?
  - b. Quelles mesures sont envisagées pour corriger les erreurs et répondre des dommages éventuels (suivi et contrôle qualité)?

---

<sup>1</sup> Li *et al.* (2019). Challenges and Pitfalls of Machine Learning Evaluation and Benchmarking. DOI : 10.48550/arXiv.1904.12437.

<sup>2</sup> Biederman and Scheirer (2020). Pitfalls in Machine Learning Research: Reexamining the Development Cycle. DOI : 10.48550/arXiv.2011.02832.