CyberSecurity

Securing Critical Business

IA et cyber-sécurité

Opportunités et challenges

Alexandre Dey, Doctorant

AIRBUS

Plan

1 **Introduction**
Les problématiques de cyber-sécurité auxquelles l'IA peut répondre

2 **Un domaine particulier**
Les spécificités de la cyber sécurité par rapports aux autres applications de l'IA

3 **Challenges communs**
Des problèmes de recherche en IA applicables en cyber-sécurité

4 **Le cas de l'adversaire**
Les techniques d'IA défensives sont vulnérables à celles offensives

5 **Conclusion**
Conclusion et travaux présents et futurs

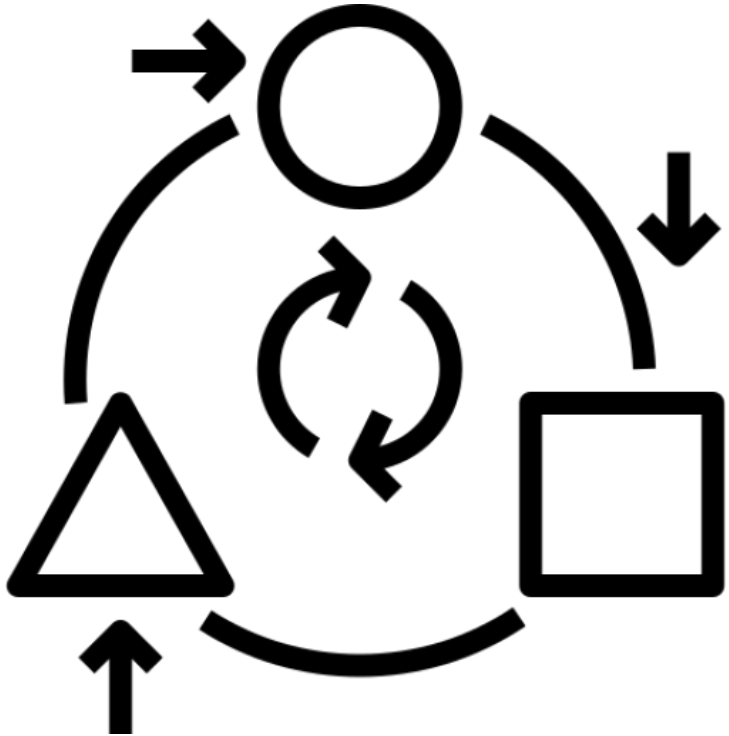
AIRBUS

CyberSecurity

Introduction

Adaptabilité

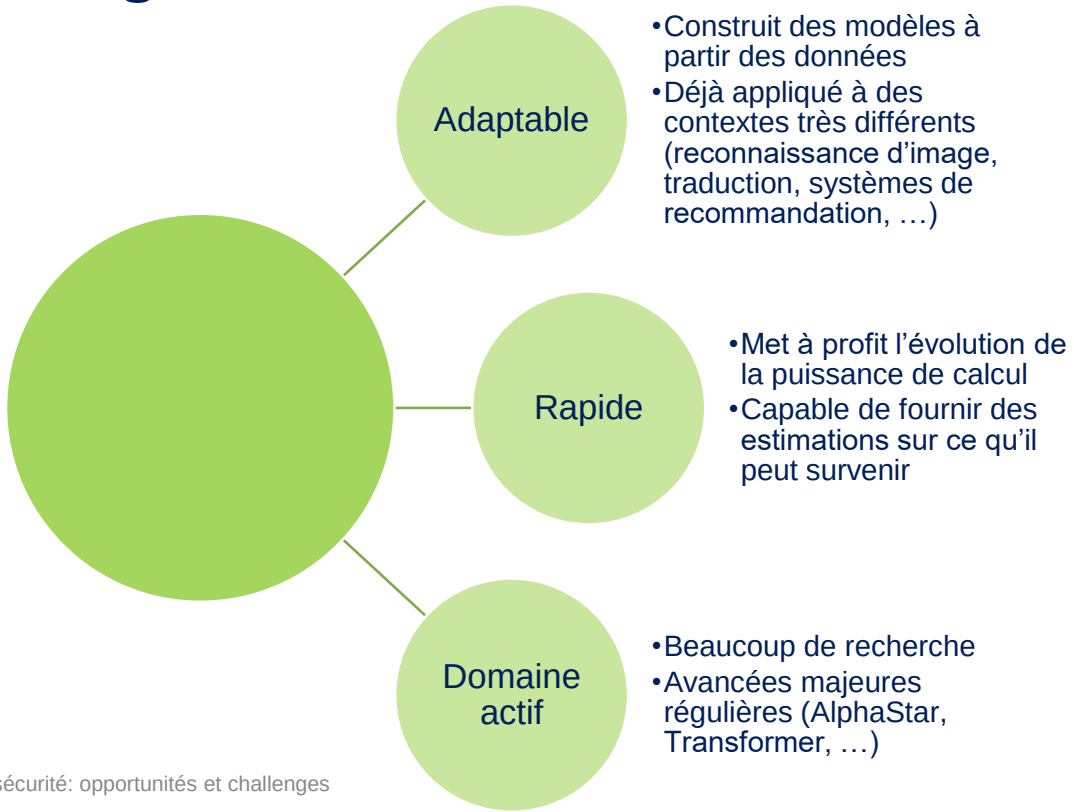
- Chaque système a ses particularités, et aussi ses cyber-menaces
- Les comportements de ces systèmes et menaces sont de plus en plus complexes
- Déployer des systèmes de détection pour un nouveau système est difficile



Rapidité

- Le volume de données à traiter est de plus en plus important
- Besoin de détecter au plus vite une intrusion pour pouvoir limiter les dégâts
- Les opérationnels ont peu de temps allouer pour traiter beaucoup de choses

L’apprentissage machine



- Construit des modèles à partir des données
- Déjà appliqué à des contextes très différents (reconnaissance d’image, traduction, systèmes de recommandation, …)

- Met à profit l’évolution de la puissance de calcul
- Capable de fournir des estimations sur ce qu’il peut survenir

- Beaucoup de recherche
- Avancées majeures régulières (AlphaStar, Transformer, …)

Un domaine particulier

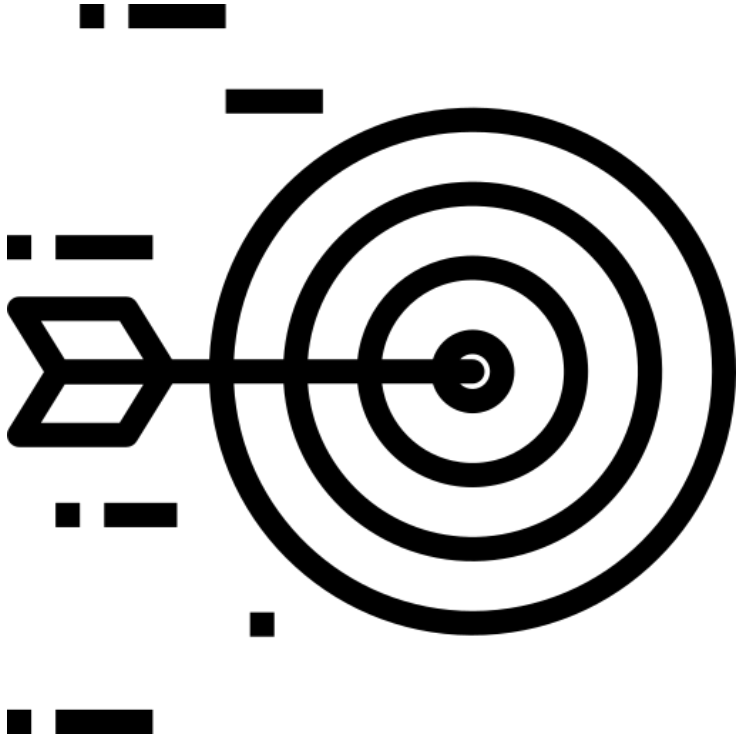
CyberSecurity

Une aiguille dans une botte de foin

- Les attaquants chercheront à être discret, à faire le moins de bruit possible sur le système cible

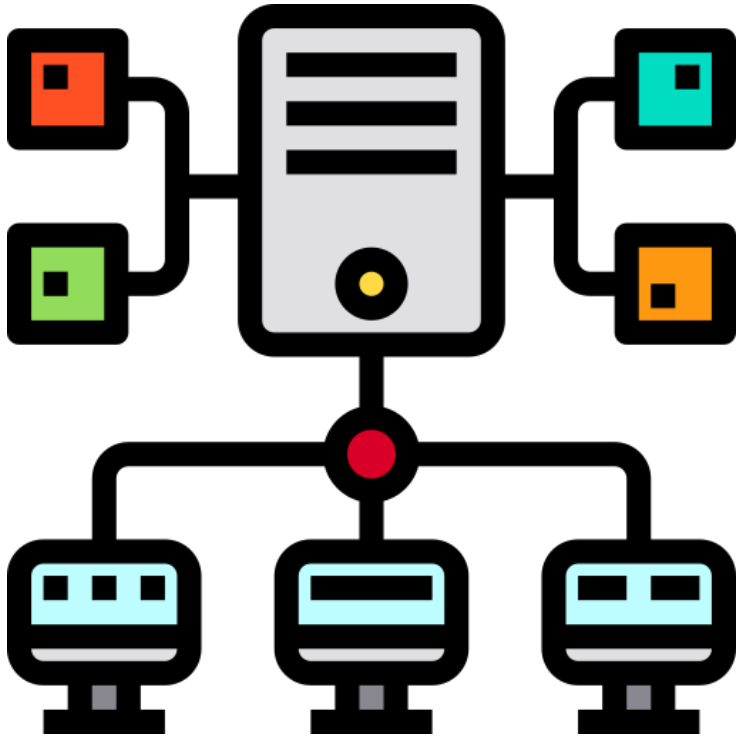
- La majorité des applications de l'apprentissage machine se concentrent sur des tendances globales, pas sur des détails

➢ **En cyber-sécurité, une anomalie peut être un indice, dans beaucoup d'autres domaines, c'est souvent du bruit**



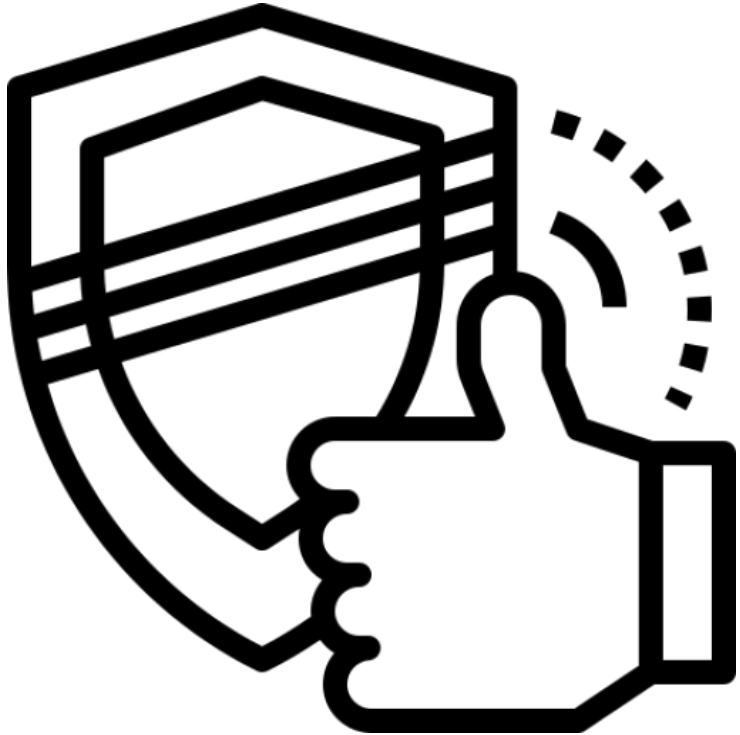
Les données

- Il est très difficile d'obtenir des données labellisées représentatives du domaine
- On peut utiliser les données de fonctionnement normal des systèmes et détecter des anomalies, mais ces données sont-elles saines ?
- D'un point de vue « data science » les données de cyber-sécurité sont complexes (mélange de nombres, de texte, de série temporelles, …)



La confiance

- Les conséquences d'une erreur peuvent être graves
- Les outils de détection sont opérés par des humains, qui doivent comprendre les résultats pour pouvoir s'y fier
- Comment peut on garantir le bon fonctionnement d'une IA ?



Challenges communs

Challenges communs

Explicabilité	• La compréhension des résultats facilite leur exploitation • Diagnostiquer les défaillances devient plus simple
Accessibilité	• Le choix des bons algorithmes et de leurs paramètres est complexe • Les opérationnels n'ont pas nécessairement d'expertise en intelligence artificielle
Challenges techniques	• Transfert de connaissance d'un système à un autre • Les comportements sur un système évoluent continuellement

Le cas de l’adversaire

Perturbation d'un classificateur

- Applicable contre certains antivirus basés sur de l'IA

CyberSecurity

Contre les détecteurs d’anomalies

- Si on chauffe progressivement la grenouille finit par accepter de se faire ébouillanter

- Une IA apprenant en continue peut être progressivement empoisonnée pour au final considérer des intrusions comme normales

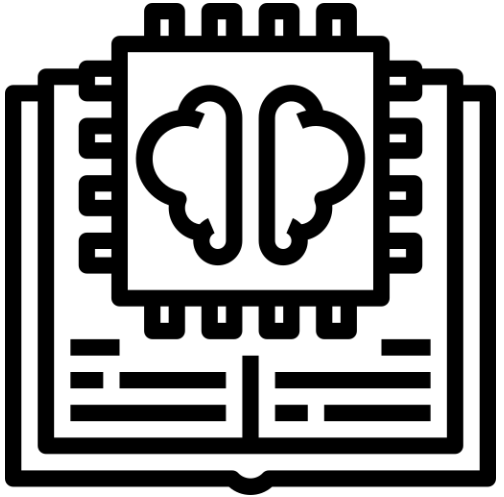
Déroulement automatique d'attaques

- Deux adversaires: un attaquant et un défenseur
- Permet de choisir automatiquement les successions d'étapes les plus susceptibles de fonctionner, les plus discrètes, les plus rapides, …
- Des premiers travaux déjà fonctionnels

Conclusion

En résumé

- L'IA peut aider à répondre à certaines problématiques en cyber sécurité
- Mais certaines spécificités compliquent sa mise en application
- L'adversaire cherchera toujours à prendre le dessus, et se servira lui aussi d'IA si nécessaire
- Beaucoup d'avancées clefs pour la cyber-sécurité seraient bénéfiques pour d'autres domaines d'application



Nos travaux

- Système de détection et de reconstruction de scénario de cyberattaques basé sur les anomalies comportementales des utilisateurs et entités

- Etude de robustesse face aux attaques type « IA adversaire »

- Usage de la CyberRange pour les tests sur des environnements contrôlés de grande ampleur


CyberSecurity
Securing Critical Business
AIRBUS