A quoi correspond lâ□□explicabilitÃ© des IA en cybersÃ©curitÃ© ?
Internet
PostÃ© par : JerryG
PubliÃ©e le : 17/1/2024 13:00:00

Lâ□□intelligence artificielle a rÃ©volutionnÃ© de nombreux domaines dont celui de la cybersÃ©curitÃ©. Cependant, cette technologie prometteuse soulÃ¨ve des questions en termes dâ□□explicabilitÃ© et de transparence. Le Machine Learning (ML)a connu des avancÃ©es remarquables depuis ces derniÃ¨res annÃ©es.

Aujourdâ□□hui, grÃ¢ce Ã dâ□□Ã©normes bases de donnÃ©es, des modÃ¨les de plus en plus Ã©laborÃ©s peuvent classifier des attaques complexes et variÃ©es sans quâ□□il soit nÃ©cessaire de les dÃ©finir explicitement.

Cependant, cette Ã©volution sâ□□accompagne dâ□□une opacitÃ© croissante. Bien que des mÃ©thodes de ML avancÃ©es, telles que les rÃ©seaux neuronaux profonds, dÃ©montrent une excellente efficacitÃ© en laboratoire, leur utilisation comme Â« boÃ®tes noires Â» peut causer des erreurs inattendues et difficiles Ã comprendre en conditions rÃ©elles.

Il est donc utile de comprendre en quoi consiste lâ□□explicabilitÃ© des IA dans le monde de la cybersÃ©curitÃ© et pourquoi cela est devenu une nÃ©cessitÃ©.

Le concept de lâ□□explicabilitÃ© des IA

Lâ□□explicabilitÃ© est la capacitÃ© dâ□□un systÃ¨me Ã rendre son processus de raisonnement et ses rÃ©sultats intelligibles pour les humains. Dans le contexte actuel, des modÃ¨les sophistiquÃ©s opÃ¨rent souvent comme des Â« boÃ®tes noires Â», masquant les dÃ©tails de leur fonctionnement.

Ce manque de transparence soulÃ¨ve des enjeux. En effet, sans une comprÃ©hension claire du processus dÃ©cisionnel, il devient difficile dâ□□identifier, et encore moins de corriger, dâ□□Ã©ventuelles erreurs. De plus, il est compliquÃ© pour lâ□□Ãªtre humain de faire confiance Ã une IA qui dÃ©livre des rÃ©sultats sans justification apparente.

Lâ□□importance de lâ□□explicabilitÃ©

Dans des domaines oÃ¹ la prise de dÃ©cision est critique, il est primordial de comprendre comment lâ□□IA opÃ¨re pour lui accorder notre confiance. Lâ□□absence dâ□□explicabilitÃ© et de transparence est aujourdâ□□hui un frein Ã lâ□□intÃ©gration de lâ□□IA dans ces secteurs sensibles.

Prenons lâ□□exemple dâ□□un analyste de sÃ©curitÃ© ; il a besoin de savoir pourquoi un comportement a Ã©tÃ© classÃ© comme suspect et dâ□□obtenir des rapports dâ□□attaque approfondis avant dâ□□engager une action aussi importante que de bloquer le trafic en provenance dâ□□adresses IP spÃ©cifiques.

Mais lâ□□explicabilitÃ© ne profite pas uniquement aux utilisateurs finaux. Pour les ingÃ©nieurs et concepteurs de systÃ¨mes dâ□□IA, elle simplifie la dÃ©tection de potentielles erreurs du modÃ¨le de ML et Ã©vite les ajustements Â« Ã lâ□□aveugle Â». Lâ□□explicabilitÃ© est donc centrale dans la conception de systÃ¨mes fiables et dignes de confiance.

Comment rendre les IA explicables

Des modÃ¨les de ML comme les arbres de dÃ©cision sont naturellement explicables. Bien quâ□□en gÃ©nÃ©ral moins efficaces que des techniques de ML plus sophistiquÃ©es comme les rÃ©seaux neuronaux profonds, ils offrent une totale transparence.

Certaines techniques â□□post hocâ□□, telles que SHAP et LIME, ont Ã©tÃ© dÃ©veloppÃ©es pour analyser et interprÃ©ter des modÃ¨les Â« boÃ®te noire Â». En modifiant les entrÃ©es et en observant les variations correspondantes dans les sorties, ces techniques permettent dâ□□analyser et de dÃ©duire le fonctionnement de nombreux modÃ¨les existants.

Lâ□□approche â□□explainability-by-design Â» va au-delÃ des techniques post hoc en intÃ©grant lâ□□explicabilitÃ© dÃ¨s la conception de systÃ¨mes dâ□□IA. PlutÃ´t que de chercher Ã expliquer les modÃ¨les a posteriori, l' Â»explainability-by-design Â» sâ□□assure que chaque Ã©tape du systÃ¨me est transparente et comprÃ©hensible. Cela peut impliquer lâ□□utilisation de mÃ©thodes hybrides et permet la conception dâ□□explications appropriÃ©es.

Lâ□□explicabilitÃ© en IA nâ□□est donc pas un luxe, mais une nÃ©cessitÃ©, surtout dans des domaines sensibles comme la cybersÃ©curitÃ©. Elle permet de gagner la confiance des utilisateurs mais aussi dâ□□amÃ©liorer continuellement les systÃ¨mes de dÃ©tection. Il sâ□□agit dâ□□un point essentiel Ã prendre en considÃ©ration dans le choix dâ□□une solution de sÃ©curitÃ©, dixit CÃ©line MINH, IngÃ©nieure en IA chez Custocy