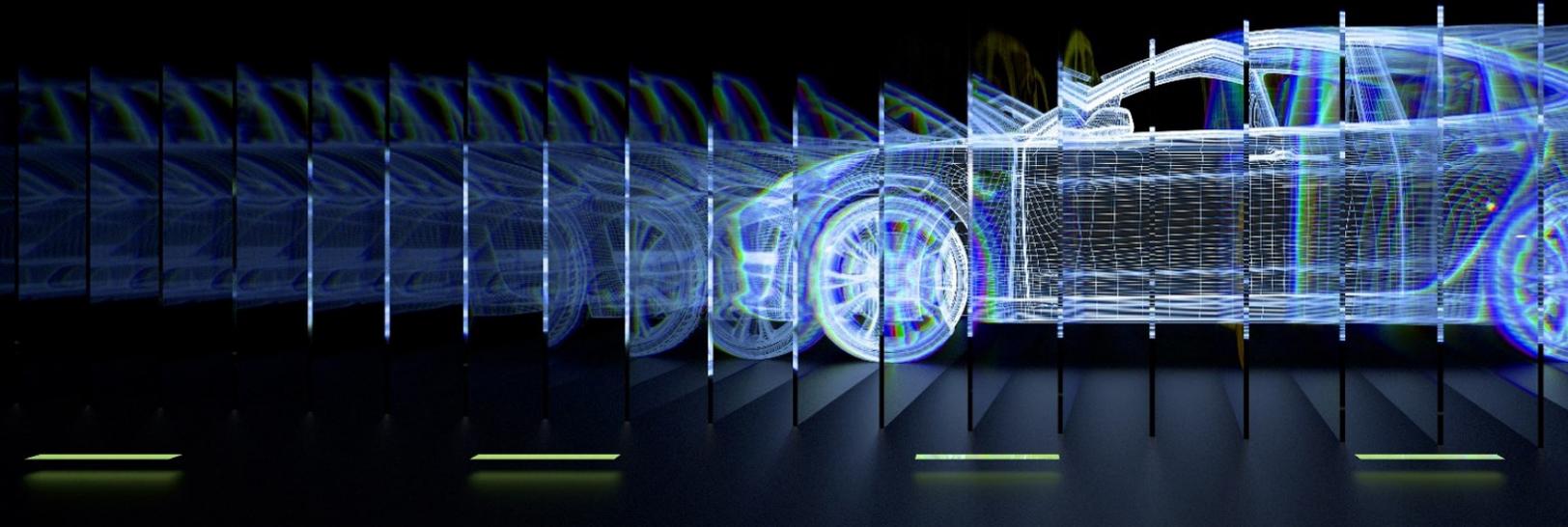


Évaluation des systèmes de perception : la précision, le rappel et la spécificité démystifiés



Date de publication : 6 septembre 2023

Résumé analytique

Le présent livre blanc explique ce que signifie la performance dans un système d'apprentissage automatique et quels sont les indicateurs de performance clés utilisés pour évaluer l'efficacité d'un modèle d'apprentissage automatique par rapport à un autre.

Dans le contexte des systèmes de perception pour les systèmes avancés d'aide à la conduite (systèmes ADAS), ce livre blanc analyse les différents indicateurs de performance clés et la manière dont ils se traduisent dans des scénarios de conduite réels. Ce document permet au lecteur de mieux comprendre les indicateurs de performance clés tels que le rappel, la précision et la spécificité et explique également des concepts tels que le vrai positif, le vrai négatif, le faux négatif, le faux positif et la matrice de confusion.

Étiquettes de classe et matrice de confusion

La sensibilité, la spécificité, la précision et le rappel sont quelques indicateurs de performance clés utilisés pour évaluer les performances d'un modèle de détection et de classification d'objets dans le cadre de l'apprentissage automatique. Avant de comprendre ces indicateurs de performance clés et la manière dont ils doivent être interprétés, il est indispensable d'assimiler les concepts d'étiquettes de classe positive et négative ainsi que de matrice de confusion.

Un vrai positif est un résultat où le modèle prédit correctement une résultante positive (c.-à-d., résultat d'un test déclaré positif à juste titre). Un vrai négatif est un résultat où le modèle prédit correctement une résultante négative (c.-à-d., résultat d'un test déclaré négatif à juste titre). Le résultat prédit est identique au résultat réel à la fois dans les instances de vrai positif et de vrai négatif. Le résultat réel est également connu sous le nom de « vérité de terrain ».

Un faux positif est un résultat où le modèle prédit à tort une résultante positive (c.-à-d., résultat d'un test déclaré par erreur positif). À l'inverse, un faux négatif est un résultat où le modèle prédit à tort une résultante négative (c.-à-d., résultat d'un test déclaré par erreur négatif). Pour comprendre ces concepts, prenons l'exemple d'une compagnie d'assurance qui utilise l'apprentissage automatique pour déterminer si un conducteur va avoir un accident et qui se sert du jeu de données ci-dessous pour former son modèle.

Tableau 1 – Exemple de jeu de données d'apprentissage

A conduit sous influence	Âge	A dépassé la limite de vitesse	A eu un accident
Oui	25	Non	Oui
Non	33	Oui	Non
...

Une fois le système formé, il est déployé pour être testé sur un nouveau jeu de données inconnu du modèle d'apprentissage automatique, mais dont les résultats réels sont connus. Par exemple, supposons que le nouveau jeu de données sur lequel le modèle est testé comporte une liste de 100 clients et que le modèle prédit que 30 clients auront un accident et que 70 clients n'en auront pas. La réalité, qui est connue, démontre que sur les 30 clients qui devaient être impliqués dans un accident, seuls dix l'ont été. De même, sur les 70 clients qui n'étaient pas censés avoir d'accident, cinq en ont eu et 65 n'en ont pas eu, comme le montre la matrice de confusion dans le Tableau 2 ci-dessous.

Tableau 2 – Matrice de confusion

		Réalité	
		A eu un accident	N'a pas eu d'accident
Prédiction	Aura un accident	10	20
	N'aura pas d'accident	5	65

Selon la matrice de confusion ci-dessus :

- Vrais positifs = 10
- Vrais négatifs = 65
- Faux positifs = 20
- Faux négatifs = 5

Sensibilité et spécificité

La sensibilité, également connue sous le nom de « rappel », est expliquée plus en détail dans la section suivante. La spécificité mesure la proportion de vrais négatifs correctement identifiés par le modèle et se calcule ainsi :

$$\text{Spécificité} = \frac{\text{Vrais négatifs}}{\text{Vrais négatifs} + \text{faux positifs}}$$

En utilisant les données de la matrice de confusion,

$$\text{Spécificité} = \frac{65}{65 + 20} = 0,76$$

L'évaluation de la spécificité est plus critique dans certains scénarios, comme la détection d'obstacles sur la route. Prenons l'exemple d'un véhicule circulant sur autoroute équipé d'un système de régulation de vitesse adaptative, de changement de voie automatique et de freinage d'urgence automatisé, toutes ces fonctionnalités étant activées. Le véhicule roule à grande vitesse lorsque des glissières de sécurité apparaissent. Certains systèmes de perception pourraient détecter à tort les glissières de sécurité comme étant un objet étranger, ce qui pourrait amener de façon inappropriée le véhicule à faire une embardée, déclencher le freinage d'urgence et/ou provoquer un accident avec d'autres véhicules sur la route. Cette situation hypothétique mais réaliste résulte d'un faux positif : la détection d'un obstacle alors qu'il n'y en a pas. L'un des objectifs d'un système de perception est de minimiser les faux positifs.

Précision et rappel en tant qu'indicateurs de performance clés

La précision est la fraction de prédictions positives correctes parmi toutes les prédictions positives faites par le modèle. En d'autres termes, la précision mesure l'exactitude du modèle dans la prédiction de l'étiquette de classe positive. Un score de précision élevé indique que le modèle prédit bien la classe positive et fait peu de prédictions faussement positives. Une précision élevée est essentielle pour améliorer la sécurité des systèmes avancés d'aide à la conduite et pour renforcer la confiance des consommateurs. Les faux positifs d'un système de perception impliquent la présence sur la route d'un objet qui n'existe pas, ce qui peut entraîner des conséquences mortelles. De tels événements provoquent un stress important chez le conducteur et les passagers du véhicule et érodent la confiance dans l'ADAS.

Mathématiquement,

$$\text{Précision} = \frac{\text{Vrais positifs}}{\text{Vrais positifs} + \text{faux positifs}}$$

En utilisant les données de la matrice de confusion,

$$\text{Précision} = \frac{10}{10 + 20} = 0,33$$

Le rappel, également connu sous le nom de « sensibilité », est la fraction de prédictions positives correctes parmi tous les positifs réels. Un score de rappel élevé indique que le modèle identifie correctement toutes les instances positives avec peu de prédictions faussement négatives. Dans les systèmes de perception et ADAS, un score de rappel élevé est essentiel. Un faible score de rappel implique un nombre élevé de faux négatifs, ce qui signifie que le système de perception est incapable de détecter un objet qui existe dans la réalité. Dans des conditions de conduite réelles, un faux négatif

peut signifier la non-détection d'une voiture, d'un piéton ou de tout autre objet, ce qui risque d'entraîner un accident.

Mathématiquement,

$$Rappel = \frac{Vrais\ positifs}{Vrais\ positifs + faux\ négatifs}$$

En utilisant les données de la matrice de confusion,

$$Rappel = \frac{10}{10 + 5} = 0,67$$

La précision et le rappel sont des mesures essentielles de la performance d'un modèle de classification et sont souvent utilisés ensemble pour évaluer l'efficacité globale du modèle. Cependant, il existe un compromis entre la précision et le rappel, car l'augmentation de l'un entraîne souvent une diminution de l'autre.

Supposons, par exemple, qu'on fixe un seuil très élevé pour qu'un modèle classe un échantillon comme positif. Dans ce cas, on peut obtenir une précision élevée mais un faible rappel parce que le modèle ne fait des prédictions positives que lorsqu'il est très confiant et peut manquer d'autres prédictions. Inversement, supposons qu'on fixe un seuil très bas pour que le modèle classe un échantillon comme positif. Dans ce cas, on peut obtenir un rappel élevé mais une faible précision parce que le modèle fait des prédictions positives même s'il n'est pas très confiant, et beaucoup de ces prédictions peuvent être incorrectes. Ces prédictions incorrectes posent un sérieux problème aux développeurs de systèmes de perception. Le système de perception doit-il être biaisé pour réduire les faux positifs et éviter les incidents dus à des objets fantômes, ou doit-il être biaisé pour réduire la proportion de faux négatifs au risque de percuter un objet? Un système de perception performant doit réduire à la fois les faux négatifs et les faux positifs.

Évaluation des solutions de fusion et de perception dans le monde réel

La fusion de capteurs est la fusion de données provenant d'au moins deux capteurs. La perception fait référence au traitement et à l'interprétation des données des capteurs afin de détecter, d'identifier, de suivre et de répertorier les objets. La fusion de capteurs et la perception permettent à un système automatisé d'aide à la conduite de développer un modèle 3D de l'environnement qui alimente l'unité de commande du véhicule.

De nombreuses solutions de fusion de capteurs font de la fusion de niveau objet, dans laquelle chaque capteur (p. ex., radar, caméra, LiDAR), avec les limitations qui lui sont propres, identifie et répertorie les objets individuellement. La technique de fusion de niveau objet exécute des algorithmes de perception basés sur l'apprentissage automatique sur des données de capteurs individuels. Cependant, cet état de choses n'est pas idéal et entraîne des performances médiocres, car aucun capteur ne peut à lui seul détecter tous les objets dans toutes les conditions, comme le montre le tableau ci-dessous, qui donne un aperçu de ses forces et faiblesses respectives. De plus, lorsque les données de capteurs ne sont pas fusionnées, le système risque de recevoir de l'information contradictoire des capteurs et de ne pas pouvoir déterminer avec un degré de certitude acceptable la marche à suivre.

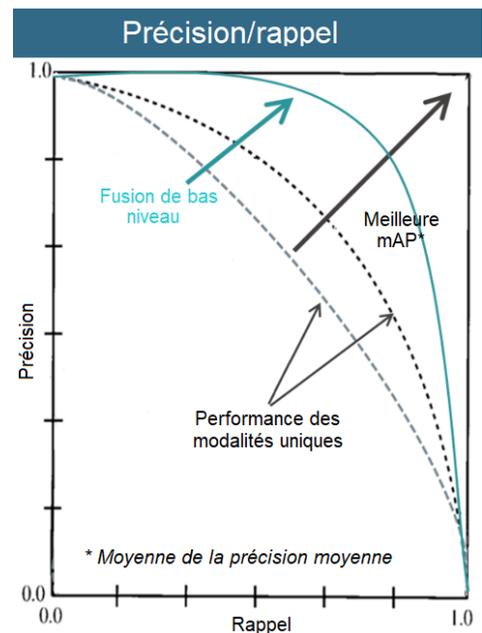
Tableau 3 – Comparaison des capteurs pour les applications ADAS et AD

	Caméra	Radar	LiDAR
Perception de la profondeur (portée)	Mauvaise	Supérieure	Bonne
Résolution	Supérieure	Mauvaise	Moyenne
Coût	Peu élevé	Peu élevé	Élevé
Performance par météo défavorable	Mauvaise	Supérieure	Bonne
Caractéristiques	Capacité de distinguer panneaux, voies de circulation, couleurs (RGB)	Longue portée, faible résolution	Portée moyenne, haute résolution

Les solutions de pointe qui fournissent un modèle environnemental 3D pour le système ADAS du véhicule s'appuient sur une fusion de bas niveau dans laquelle les données brutes de tous les capteurs à bord d'une voiture sont fusionnées (compensant ainsi les faiblesses des autres capteurs), puis des algorithmes d'apprentissage automatique visant la perception sont exécutés sur ce jeu de données combiné.

LeddarVision™ est une plateforme bas niveau de fusion de capteurs et de perception qui génère un modèle environnemental 3D intégré multicapteurs pour configurations caméra, radar et LiDAR dédiée aux systèmes avancés d'aide à la conduite. Les principales caractéristiques de LeddarVision sont ses performances élevées, son coût, sa modularité et son indépendance des capteurs. Toutefois, dans le cadre de ce livre blanc, l'accent sera mis sur l'évaluation des performances au regard des paramètres examinés dans les sections précédentes.

Idéalement, un modèle d'apprentissage automatique devrait avoir des scores de précision et de rappel égaux à un, c.-à-d., zéro faux positif ou négatif, cela restant l'objectif, à savoir construire des modèles qui offrent des performances aussi proches que possible de un. Alors que les techniques traditionnelles, telles que la fusion de niveau objet, proposent un compromis entre précision et rappel et se limitent à fournir des performances comprises dans la ligne pointillée noire, la fusion de bas niveau ne se contente pas d'un tel compromis mais repousse les limites du progrès technologique en fournissant une meilleure moyenne de la précision moyenne (mAP¹). Ce dernier concept est une mesure utilisée pour évaluer les modèles de détection d'objets tels, notamment, que Fast R-CNN, YOLO² et Mask R-CNN. La technologie de fusion des données brutes utilisée permet de détecter des obstacles de très petites dimensions sur la route avec un meilleur taux de détection et moins de fausses alertes que les solutions existantes axées sur la fusion de données de niveau objet.



¹ Mean average precision.

² "You Only Look Once".

Résumé

Les fausses alertes sont un élément essentiel des systèmes de fusion de capteurs et de perception et ont des conséquences importantes sur les performances des systèmes ADAS, conséquences allant de l'activation intempestive et déplaisante d'un système d'alerte à la collision mortelle. Cependant, même l'activation accidentelle et apparemment anodine d'un système d'alerte peut avoir de graves conséquences.

L'une des façons d'évaluer les modèles d'apprentissage automatique consiste à utiliser des indicateurs de performance clés tels que la spécificité, la précision et le rappel (c.-à-d., la sensibilité), en attribuant un nombre mathématique à la performance du modèle. Tandis que la technologie existante de fusion de niveau objet permet seulement de jouer sur la précision et le rappel, la fusion de bas niveau repousse les limites de la performance pour fournir une « moyenne de la précision moyenne » supérieure, une mesure utilisée pour évaluer la performance.

Consultez les ressources suivantes pour en savoir plus sur la fusion de bas niveau, la technologie et les produits proposés par LeddarTech, ainsi que les ressources techniques et industrielles.

[Vidéos de démonstration](#)

[LeddarVision](#)

[Produits pour niveaux d'autonomie 2/2+](#)

[FAQ sur la perception](#) (en anglais)

[Les réseaux neuronaux profonds dans la perception](#)

[Balado sur la fusion de bas niveau](#)

Le présent livre blanc ne constitue pas un modèle de référence. Les recommandations contenues aux présentes sont fournies « en l'état » et sans garantie quant à leur exhaustivité ou leur exactitude.

LeddarTech® a tout mis en œuvre pour s'assurer que les renseignements contenus dans le présent document sont exacts. La totalité des renseignements contenus aux présentes sont fournis « en l'état ». LeddarTech ne pourra être tenue pour responsable d'aucune erreur ou omission dans le présent document ni d'aucun préjudice découlant de l'information contenue aux présentes ou y afférent. LeddarTech se réserve le droit de modifier la conception ou les caractéristiques de ses produits à tout moment, sans préavis et à sa seule discrétion.

LeddarTech ne répond pas de l'installation de ses produits ni de l'usage qui en est fait, et décline toute responsabilité si un produit est utilisé pour une application pour laquelle il ne convient pas. Il vous incombe entièrement (1) de sélectionner les produits appropriés pour votre application, (2) de valider, concevoir et tester votre application, et (3) de vous assurer que votre application répond aux normes de sûreté et de sécurité en vigueur.

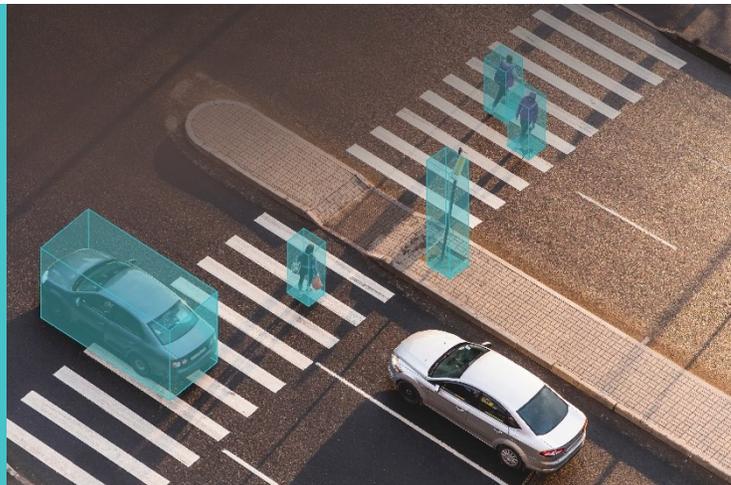
De plus, les produits LeddarTech sont assujettis aux conditions générales de vente de LeddarTech ou autres conditions applicables convenues par écrit. En achetant un produit LeddarTech, vous vous engagez également à lire attentivement l'information contenue dans le guide d'utilisation qui accompagne le produit acheté et à y être lié.

Leddar, LeddarTech, LeddarVision, LeddarSP, VAYADrive, VayaVision et les logos associés sont des marques de commerce ou des marques déposées de LeddarTech Holdings Inc. et de ses filiales. Tous les autres noms de marques, noms de produits et marques sont ou peuvent être des marques de commerce ou des marques déposées utilisées pour désigner les produits ou les services de leurs propriétaires respectifs.

À propos de LeddarTech

Entreprise mondiale de logiciels fondée en 2007, basée à Québec et disposant de centres de R&D supplémentaires à Montréal, Toronto et Tel Aviv (Israël), LeddarTech développe et propose des solutions logicielles complètes de fusion bas niveau de capteurs et de perception reposant sur l'intelligence artificielle qui permettent le déploiement d'applications ADAS, de conduite autonome (AD) et de stationnement. Le logiciel de classe automobile de LeddarTech applique des algorithmes d'intelligence artificielle et de vision numérique avancés afin de générer des modèles 3D précis de l'environnement, pour une meilleure prise de décision et une navigation plus sûre. Cette technologie performante, évolutive et économique permet la mise en œuvre efficace de solutions ADAS pour véhicules automobiles et hors route par les équipementiers et les fournisseurs de rang 1 et 2. Ayant déposé plus de 150 demandes de brevets (dont 80 accordées) qui améliorent les capacités des systèmes d'aide à la conduite, de conduite autonome et de stationnement, l'entreprise a contribué à plusieurs innovations liées à des applications de télédétection. Une meilleure perception sensorielle de l'environnement autour du véhicule est essentielle pour rendre la mobilité plus sûre, plus efficace, plus durable et plus abordable : c'est ce qui motive LeddarTech à vouloir devenir la solution logicielle de fusion de capteurs et de perception la plus largement adoptée.

Renseignements complémentaires : sales@leddartech.com



LeddarTech®

CANADA – ÉTATS-UNIS – AUTRICHE – FRANCE – ALLEMAGNE – ITALIE – ISRAËL – HONG KONG – CHINE

Siège social

4535, boulevard Wilfrid-Hamel, bureau 240
Québec (Québec) G1P 2J7, Canada

leddartech.com

Tél. : + 1-418-653-9000

Sans frais : 1-855-865-9900