



Criteo annonce le jeu de données le plus important de l'industrie pour l'apprentissage statistique à l'attention de la communauté des chercheurs

Plus d'un téraoctet de données sera mis à la disposition des chercheurs pour évaluer les algorithmes d'apprentissage distribués dans le domaine de la recherche

Paris, France – le 18 juin 2015 - [Criteo](#) (NASDAQ:[CRTO](#)), la société technologique spécialiste du Marketing à la Performance, annonce la disponibilité du plus important jeu de données d'apprentissage statistique (*machine learning*) au service de la communauté open source. L'objectif de Criteo est de soutenir la recherche académique et l'innovation dans le domaine des algorithmes d'apprentissage statistique.

Face à l'importance croissante des problèmes liés aux données à grande échelle dans différents secteurs dont la publicité à la performance, la mise à disposition de jeux de données d'une telle ampleur est nécessaire pour favoriser les avancées industrielles. Les jeux de données anonymes extraits d'applications concrètes permettent aux chercheurs de tester, d'affiner et d'améliorer les plateformes d'apprentissage statistique auxquelles un grand nombre d'entreprises recourent. À titre d'exemple, Criteo s'appuie sur ses propres algorithmes d'apprentissage pour prévoir avec précision à quel moment un consommateur est le plus enclin à cliquer sur une annonce, augmentant ainsi le retour sur investissement des annonceurs et de leurs campagnes.

« *La précision et la rapidité des algorithmes d'apprentissage statistique sont deux éléments indispensables au succès de notre entreprise et de nombreuses autres, mais ils seraient irréalisables sans la possibilité d'accéder à des jeux de données publiques.* » affirme **Olivier Chapelle, Principal Research Scientist chez Criteo**. « *En tant que participant actif à de nombreux projets open source, Criteo s'engage à soutenir les professionnels de l'apprentissage statistique pour contribuer à l'innovation open source. Conscients du rôle essentiel que jouent les chercheurs indépendants, nous avons rendu disponible ce jeu de données unique pour faciliter l'innovation, au bénéfice de la communauté académique et de l'industrie au sens large, dans le cadre d'une collaboration visant à améliorer les technologies d'apprentissage statistique.* »

Criteo est en mesure de fournir à la communauté des chercheurs indépendants un jeu de données unique grâce à son envergure hors pair, aux relations directes nouées avec des éditeurs et des annonceurs, ainsi qu'au haut niveau d'engagement des consommateurs avec les publicités à la performance que la société génère. Criteo reçoit 30 milliards de requêtes HTTP par jour (jusqu'à deux millions par seconde), affiche trois milliards de bannières publicitaires uniques par jour, et stocke quotidiennement 20 téraoctets de données nouvelles avec une capacité de stockage brut de 37 pétaoctets.

Avec plus de quatre milliards de lignes totalisant plus d'un téraoctet, ce nouveau jeu de données s'appuie sur le jeu de données de prédiction des clics de Criteo, initialement lancé dans le cadre de son « [Display Advertising Challenge](#) » mené avec Kaggle. Le jeu de données a déjà été utilisé [en tant que plateforme de test](#) par des chercheurs de l'Université Carnegie Mellon ; de nouvelles applications universitaires et de recherche sont prévues.

« *Le jeu de données de 1 To de Criteo nous a été très précieux pour évaluer l'évolutivité des algorithmes d'apprentissage utilisés dans l'estimation des taux de clics à haut débit que nous développons dans le cadre de notre projet Marianas Labs* », a déclaré **Alexander Smola, professeur à l'Université Carnegie Mellon**, Pittsburgh (Pennsylvanie). « *Nous sommes impatients de poursuivre notre collaboration avec Criteo sur des jeux de données supplémentaires.* »

Le téraoctet de données de Criteo est hébergé en environnement Microsoft Azure ; les détails concernant la façon d'y accéder, de l'utiliser et de le télécharger sont disponibles sur le site [Criteo Labs](#). Pour plus d'information sur les solutions de publicité numérique à la performance de Criteo, vous pouvez visiter www.criteo.com.

###

À propos de Criteo

Criteo est une société mondiale, leader de la publicité numérique à la performance, avec plus de 7 800 sociétés clientes dans le monde. Criteo est présent dans plus de 70 pays avec plus de 1 500 salariés présents dans des bureaux répartis à travers les États-Unis, l'Europe et l'Asie. Pour plus d'information, rendez-vous sur <http://www.criteo.com>.

###

Contacts presse

Alexandra Pelissero
Criteo
Tél : +331 70 56 79 12
A.pelissero@criteo.com

Sabrina Gonzalez, Sébastien Meunier, Morad Salehi
Hotwire PR
+331 43 12 55 74 / 77 /65
CriteoFR@HotwirePR.com