



Cyrès, à la pointe du Data Mining, avec Apache Spark

Paris, - Ingensi, business unit du groupe Cyrès, précurseur français dans l'exploitation et la valorisation des données des entreprises, poursuit son effort d'innovation et met à la disposition de ses clients son expertise sur les solutions Big Data les plus pointues du marché.

Suivre l'évolution naturelle d'Hadoop

Avec l'arrivée du Big Data et du framework Hadoop, la technologie datamining a fait apparaître de nouvelles valeurs autour de l'information détenue en entreprise, par le biais du prédictif. Inépuisable terrain de recherche, les technologies Big Data profitent maintenant de vitesses de traitements surprenantes grâce à Apache Spark, qui améliore sensiblement l'application des algorithmes de datamining. Par ailleurs plus performante, la technologie Spark s'intègre avec les autres briques de l'écosystème Hadoop et propose, compte tenu de ses performances, des fonctionnalités très attendues par les experts, comme le traitement des données en temps réel.

Pour répondre à cette demande croissante d'expertise sur l'exploitation des données et l'élaboration de modèles prédictifs utiles au business des entreprises, Cyrès additionne les innovations en continu issues de sa cellule R&D à l'expérience de ses ingénieurs Big Data sur Hadoop. C'est dès le début de l'aventure Spark que les équipes d'Ingensi ont suivi, testé et adopté le framework, avec en toile de fond la croyance, aujourd'hui vérifiée, qu'il serait irrémédiablement le successeur de Map Reduce. Très vite les équipes d'Ingensi ont su évaluer les nouveaux possibles à offrir à leurs clients, notamment sous l'angle du machine learning auquel Spark se prête admirablement. Classification, regroupement et prédiction connaissent un nouveau souffle.

Une valeur fondée sur le diptyque Hadoop / Spark

Ingensi a choisi de déployer Spark sur son cluster Hadoop. La grande force d'hadoop est de pouvoir construire des référentiels en entreprise et simplifier la gestion de l'information en la proposant dans un espace partagé, réunissant à la fois les données et les outils performants pour les traiter. Même si le cœur d'Hadoop connaît aujourd'hui une vraie révolution avec Spark (à tel point que certains semblent choisir de s'en passer), hadoop demeure une des rares solutions performantes pour répondre aux besoins de business intelligence, particulièrement en association avec Spark (ou Impala).

Une offre de service complète

Concrètement, Ingensi met à disposition de ses clients les outils et l'infrastructure nécessaires au chargement des données et au traitement. Cette étape de traitement définit le use case qui déclenchera les travaux de recherche, visant à répondre à un besoin fonctionnel détecté en entreprise. A ce stade, les équipes d'Ingensi travaillent en étroite collaboration avec le client au travers d'ateliers et d'audits. C'est une étape de conseil cruciale dans le déroulement du projet.

Les objectifs à atteindre ainsi définis, la troisième phase s'articule autour de la définition d'un prototype afin d'adresser le use case, complété d'un rendu de résultats probants et de la démonstration d'un retour sur investissement. Validé, le prototype passe en production et la capacité de la plateforme est augmentée afin d'accueillir autant de données que de besoin.

Se qualifier sur le Big Data

C'est précisément cette démarche qui a été adoptée récemment avec une grande société d'assurance dont la problématique portait sur la capacité de déterminer avec justesse les taux de réussite ou d'échec dans la transformation d'un devis.

Les données anonymisées ont été traitées à partir du duo technologique Hadoop / Spark (via sa librairie MLlib), capable d'une proposition de valeur supérieure et adaptée aux entreprises en démarrage ou en voie de qualification sur le Big Data.

À propos de Cyrès.

[Visualiser l'article](#)

Le groupe Cyres est une société technologique qui depuis 15 années administre le patrimoine numérique de plus de 1000 entreprises et ce autour de deux business units : Ingensi et Hosteam.

Ingensi, précurseur français dans le traitement et l'analyse du Big Data est notamment à l'origine d'Analytic Suite Appliance, une plateforme applicative destinée à fédérer l'ensemble des informations internes et externes au sein d'un référentiel unique

Hosteam, expert dans le déploiement d'infrastructures matérielles hautement sécurisées, offre l'accès à un Cloud de proximité français et modulaire. Hosteam délivre des solutions de Cloud privé, Cloud hybride, Cloud Big Data, adossées à une gamme de services à forte valeur ajoutée. Hosteam permet aux entreprises d'héberger des applications métiers, solutions Big Data, bases de données, solutions collaboratives ou solutions de virtualisation de poste de travail, le tout dans un Cloud français et ultra sécurisé.

Basé à Tours et Paris le groupe affiche une croissance moyenne de 14 %.

Pour en savoir plus :

<http://www.cyres.fr/> <http://www.cyres.fr/cyres/ingensi> <http://www.cyres.fr/cyres/hosteam>