



Mise en place d'une démarche DataOps



Contexte

La MAIF a ouvert son Datalab à la fin de l'année 2017 pour mener à bien des projets de DataScience et d'innovation. Le Datalab a voulu mettre l'accent sur la production et l'industrialisation de ses travaux, afin de tirer une réelle valeur de ses projets et éviter l'écueil du « Shadow IT » dans les entités qu'elle doit servir.

Saegus a ainsi été sollicité pour cadrer et réaliser ces processus de mise en production et d'industrialisation, mais aussi pour accompagner et acculturer les entités utilisatrices à ces nouvelles méthodes.

Nous avons ainsi pu travailler d'une manière macro en définissant les process, les normes et bonnes pratiques, mais aussi à des niveaux projets, plus micro, où nous avons développé et réalisé les mises en production de modèles prédictifs et de flux de transformation de données.

Focus projet email

La MAIF reçoit quotidiennement des dizaines de milliers de mails de ses sociétaires dont les délais de réponse peuvent être longs. Elle souhaite ainsi réduire ces délais d'attente pour rester en phase avec sa volonté d'excellence dans le service client.

Pour mieux router les emails, la MAIF a développé un modèle prédictif, associant NLP et Deep Learning, qui classe les emails dans des catégories. Nos équipes ont mis en production ce modèle, tout en mettant en place sa méthodologie DataOps.

Réalisations

Pour répondre à ce type de problématique, nous avons réalisé les actions suivantes :

- Création d'un cadre de référence technique (sélection et recommandation d'assemblages technologiques en fonction des uses cases) ;
- Mise en place des normes et bonnes pratiques de développement pour les Data Engineers et Data Scientists ;
- Création de Starter Kit pour les différentes technologies (Starter Pyspark, Starter Python, Starter API prédictive ...) ;
- Coaching des Data Scientist sur les projets pour leur appropriation de ces nouvelles normes ;
- Déploiement d'un environnement et des pratiques « Data Ops » sur tous les projets Data ;
- Acculturation et adoption des nouveaux processus aux différentes entités ;
- Utilisation de Git dans les projets.

Focus projet email

Pour le cas particulier du projet, nous avons effectué différentes tâches :

- Encapsulation du modèle dans une API REST asynchrone (Flask) ;
- Dockerisation de l'API ;
- Intégration continue avec tests unitaires ;
- Déploiement continu sur une plateforme d'orchestration de containers (Rancher) ;
- Suivi et monitoring de l'API (Prometheus).

Processus de mise en production

Meilleure gestion des versions (Git)

Cycles de vies raccourcis (intégration et déploiement continus)

Rapidité au démarrage des projets grâce aux Starter Kits

Focus projet email

- Mise à l'échelle possible et simple
- Evolutions déployables rapidement
- Performances élevées

Facteurs clés de succès



Une appropriation des problématiques fonctionnelles



Une démarche pédagogique auprès des utilisateurs



Une synergie agile mise en place entre l'équipe de développeurs et l'équipe socle



Triple compétence technique Engineering, DataScience & DevOps



Compréhension des exigences de chacun



Usage des outils déjà présents à la MAIF



Frédéric de Javel
Chef de projet
MAIF

« Avec la multiplication des moyens digitaux, nous recevons des milliers de mails par jour ! À l'aide d'algorithmes de traitement automatique du langage et de deep learning, nous avons optimisé leur routage. Saegus nous a permis d'être agile tout au long du cycle de vie de ce projet DataScience en structurant notre ingénierie de développement, en encapsulant le modèle dans une api dockerisée et en implémentant l'intégration et déploiement continu. L'API prédictive tourne aujourd'hui en production avec des performances élevées. Ses évolutions sont désormais facilitées et ses temps de déploiements nettement plus rapides grâce à Saegus et leur expertise DataOps. »

Vos contacts chez Saegus



Frédéric BRAJON
Co-fondateur et Directeur Général
frederic.brajon@saegus.com

Nicolas RISI
Référént DataOps
nicolas.risi@saegus.com

