AXYS Consultants

Exercice 2020

Projets d’Innovation

(Crédit d’Impôt en faveur de l’Innovation)

**SOMMAIRE**

1. [PRÉSENTATION DE L’ENTREPRISE 3](#_Toc91499462)

[1 L’entreprise 3](#_Toc91499463)

[2 Gestion de l’innovation 3](#_Toc91499464)

1. [SYNTHÈSE TECHNIQUE 4](#_Toc91499465)

[1 Développement d’outils technologiques spécialisés 4](#_Toc91499466)

[1.1 Objets du projet 4](#_Toc91499467)

[1.2 Contexte économique du projet 5](#_Toc91499468)

[1.3 Description des travaux effectués 6](#_Toc91499469)

[1.4 Conclusions 16](#_Toc91499470)

1. [ANNEXES 18](#_Toc91499471)

# PRÉSENTATION DE L’ENTREPRISE

## L’entreprise

Axys Consultants, société de conseil en management fondée en 1987 et comptant 163 consultants, accompagne ses clients, grands groupes et ETI, sur leurs problématiques de transformation dans les domaines des achats, de la finance, du marketing digital et du commerce.

Axys Consultant place l’usage de la donnée comme catalyseur de la conduite du changement notamment à travers ses deux BU transverses, Data et Change.

En effet, la mutation actuelle du marché du conseil et l’émergence de nouveaux modèles disruptifs et l’accroissement de problématiques basés sur l’IA, la Blockchain et la RPA, transformant la chaîne de valeur, incitent les cabinets à repenser la manière dont ils conseillent leurs clients.

Afin de se positionner en tant que leader sur ces problématiques, la société a ouvert son LAB IA début 2019. L’objectif est de démontrer les avantages offerts par le Machine Learning pour ainsi accroitre la valeur ajoutée de ses prestations grâce à l’IA « As a Service ».

De cette manière, les travaux de recherche menés sur ces sujets nous permettent d’intégrer des prestations dans domaines à forte valeur ajoutée à notre offre de services, tout en ayant la capacité de doter nos consultants de nouveaux outils dans le but d’améliorer notre productivité et la qualité de nos prestations.

## Gestion de l’innovation

Afin de relever ces défis, le LAB IA d’Axys Consultants peut s’appuyer sur l’expertise et l’expérience de ses équipes. En effet, celle-ci est constitué d’ingénieurs développeurs, de Data Scientists et Data Analysts bénéficiant tous d’une expérience significative dans le secteur de la tech.

De plus, nos travaux font régulièrement l’objet de publications sous forme de tribunes ou de livres blanc sur notre page internet ainsi que sur des sites spécialisés.

Par ailleurs, des travaux d’innovation sont également menés par les équipes Axys Consultants qui travaillent notamment au développement de nouvelles solutions technologiques à destination de ses clients.

# SYNTHÈSE TECHNIQUE

**Chefs de projet : M. Jean-Luc Marini**

**M. Lionel Bianchi**

**M. Marc Mironneau**

**M. Julien Samarcq**

## Développement d’outils technologiques spécialisés

### Objets du projet

Durant l’année 2020, Axys Consultant a travaillé sur diverses solutions technologiques afin de proposer des solutions innovantes et adaptées aux besoins de ses clients exerçant dans de multiples secteurs d’activités.

Dans un premier temps, nous avons collaboré avec notre client **BCA Expertise**, une entreprise d’expertise automobile qui intervient sur toute la nature des dommages qu’un véhicule a subi lors d’un accident. Nous avons développé une application Blockchain permettant d’optimiser le processus de gestion des véhicules gravement accidentés. Cette plateforme permet aux différentes parties prenantes de collaborer plus efficacement via un ordonnancement des actions et une identification des acteurs concernés à chacune des étapes tout en consignant l’ensemble des décisions et des actions effectuées au cours du processus.

En outre, nous étions intervenus dans l’entité IES de la société BCA qui opère plusieurs micro-services internes basés sur l’Intelligence Artificielle dans le domaine de l’assurance automobile notamment, les services de sécurité, supervision, gestion des clients, stockage des données, Backup, service Mesh, développements IA ainsi que le développement des logiciels.

Si aujourd’hui les solutions IA de l’IES rendent les services attendus, ils n’offrent pas les meilleures garanties en termes de scalabilité, de gestion multi-clients, de supervision et de bonnes pratiques de déploiement. De plus, l’interaction entre les différents services n’est pas conçue de façon optimale. Nous avons donc développé une architecture Micro-Services SaaS dont l’objectif est de supporter la charge induite par le nouveau positionnement de l’IES en tant que fournisseur de services, permettant d'intégrer rapidement de nouvelles offres de services et fournir tous les outils nécessaires au suivi de la production, tout en limitant au maximum l'impact de ce changement d'architecture sur les micro-services eux-mêmes.

De plus, nous avons collaboré avec l’opérateur télécom **La Poste Mobile** qui a souhaité développer un algorithme multicritère de gestion de la fragilité financière permettant d’identifier et de quantifier le risque de défaut de paiement des clients lors de la souscription pour un opérateur télécom. L’objectif est ainsi de diviser par deux le taux de résiliation des abonnements à trois mois tout en limitant les demandes de garanties aux clients les plus à risque.

Nous avons enfin développé, pour notre client **SKODA,** un outil de prédiction et pilotage des leads permettant de prédire le nombre de leads qualifiés dans le secteur automobile en fonction du plan marketing (publicité, messages, budget par partenaire, etc.,) ainsi que des données externes.

Les travaux menés en 2020 s’inscrivent dans la continuité des initiatives de recherche entreprises depuis le lancement du projet dans le but d’offrir une plateforme toujours plus innovante et performante à nos utilisateurs.

### Contexte économique du projet

Aujourd’hui, de plus en plus d’entreprises fournissant des solutions blockchain dans les différents secteurs d’activités voient le jour. IBM constitue un acteur majeur dans ce domaine fournissant différentes solutions d’approvisionnement en blockchain notamment, les solution Trust Food et Trade Lens qui sont deux projets de très grande envergure réussis par IBM.

Trust Food est un réseau mettant en relation les acteurs de l’approvisionnement alimentaire par le biais d’un enregistrement partagé et permissionné des données du système alimentaire, tandis que Trade Lens est une plateforme de chaîne d’approvisionnement permettant le partage d’informations et la collaboration entre les chaînes d’approvisionnement tout en réduisant les frictions commerciales afin de favoriser le commerce mondial.

L’ensemble des compétiteurs sont limités à des solutions partielles, applicables à un cas de figure précis. A notre connaissance, nous sommes les premiers à offrir une plateforme basée sur la blockchain dans le secteur automobile permettant d’optimiser le processus de gestion des véhicules gravement accidentés.

Concernant notre solution de détection de fragilité financière, notre étude du marché nous a permis de constater qu’il n’existe actuellement aucune solution adressant ce besoin de prédire le taux de résiliation des abonnements pour les opérateurs télécom. Pour leurs contrôles, les opérateurs télécom ont l'habitude d'interroger la base Preventel (résiliation autres opérateurs) et d'analyser la conformité des données bancaires. Les solutions existantes permettent la détection de fraudes financières dans des cas de figures différents au nôtre. Nous pouvons citer à titre d’exemple, la solution **BankWare.NET** qui est un logiciel de détection de fraude financière dédiée aux banques.

Enfin, en ce qui concerne notre solution de prédiction et pilotage des leads, nous avons constaté qu’il existe plusieurs solutions puissantes dans le marché comme les solutions TilKee et SalesForce. Tilkee est un logiciel permettant de mieux gérer les propositions commerciales afin d’augmenter les probabilités de conclure une vente tandis que SalesForce est une solution CRM permettant de gérer les leads et les opportunités de ventes. Cependant, ces solutions se limitent à mesurer la contribution de chaque levier sans prise en compte des éléments externes. En contrepartie, notre modèle permet d'intégrer des données externes en prenant en compte l'impact des différentes campagnes commerciales (ex. campagne en ligne vs campagne TV vs campagne radio).

Sur le plan du pilotage, nous nous distinguons par rapport aux solutions existantes par le choix des KPI et l'ergonomie de la solution. Celle-ci a reçu le prix du Data Festival en qualité de meilleur dispositif de pilotage.

### Description des travaux effectués

En 2020, nous avons travaillé sur diverses solutions technologiques afin de proposer des solutions innovantes et adaptées aux besoins de ses clients agissant dans différents secteurs d’activités.

#### Blockchain

L'objectif du projet consiste à développer une application pour notre client BCA Expertise, une entreprise d’expertise automobiles qui intervient sur toute la nature des dommages qu’un véhicule à subit lors d’un accident. L’application permet de faciliter le processus de gestion des véhicules gravement accidentés en faisant collaborer les différentes parties prenantes (Garage, Assurance, Experts, Préfecture) grâce à une plateforme technique basée sur la technologie "Blockchain".

Le but est alors de créer un réseau d’acteurs pour diminuer considérablement les frais et les délais d’expertise de transmission de documents, avoir des certifications immuables pour éviter les fraudes, avoir un registre commun et transparent (visible par tous les acteurs de la blockchain) et aussi réduire les coûts et le travail sans valeur ajoutée.

Contrairement à la situation actuelle où les échanges entre les différents acteurs du processus sont réalisés indépendamment et de façon non structurée, l'application permet d'ordonnancer les actions de chacun des acteurs et de tracer dans une blockchain comme dans un livre de comptes les décisions prises et les actions effectuées à chaque étape du processus.

Ce type de projet nécessite de base une infrastructure solide en matière de sécurité, de robustesse et de scalabilité. Mais au-delà de ces considérations habituelles, ce projet nécessitait de mettre en œuvre une blockchain qui n'impose pas aux différents acteurs une infrastructure technique (datacenter, cloud) particulière. En effet, afin d'obtenir l'adhésion de tous les acteurs, il faut pouvoir leur laisser la possibilité de contribuer à la blockchain dans l'environnement qu'ils choisissent.

##### Configuration de la console

Tout d’abord, nous avons de configuré la console IBM blockchain via le cloud IBM qui propose une multitude de fonctionnalités liées aux différents outils proposés. Cette console permet d’ajouter les différents acteurs, les différents privilèges ainsi que les possibilités de transactions.

Pour créer notre propre blockchain, nous avons commencé par créer un service Kubernetes. C'est un conteneur d’applications qui permet un déploiement automatisé dans le cloud IBM, ce qui nous permet de déployer un cluster qui fait office de base de données distribuée. Ensuite, nous avons lancé la plateforme pour décider du consortium, soit les participants du réseau (Voir Figure 1).

Dans notre cas, nous avons défini 3 acteurs notamment : le ministère, les experts automobiles et les compagnies d’assurances. Une fois que le nous avons décidé des composants comme le nombre d’organisations présentes, le nombre d’acteurs présents dans chacune d'elles, nous avons affecté «l’ordering Node» et les certificats d’autorités. (Voir Figure 2)

Une image contenant texte, moniteur, capture d’écran, écran

Description générée automatiquement

Figure 1 : Différents acteurs de la blockchain

Une image contenant texte, moniteur, mur, capture d’écran

Description générée automatiquement

Figure 2 : Affectation des certificats d’autorité

De nombreuses blockchains distribuées n’ont pas ces attributs précédemment cités. C’est-à-dire qu'elles fonctionnent avec des processus de consensus basés sur des algorithmes de probabilités complexes qui garantissent la cohérence des processus par rapport au registre commun à un degré élevé, par exemple Bitcoin ou Ethereum. En revanche l’outil de conception de blockchains privées Hyperledger est composé d’un nœud différent des autres nommé « *ordering Node*». C’est celui qui gère toutes les transactions et s’assure de leur conformité au sein du réseau. Les certificats d’autorités sont les permissions accordées à certains utilisateurs dans chaque organisation, ce sont les Membership Service Provider (MSP).

Les organisations sont les différentes parties prenantes. Dans le cas de notre PoC, les différentes parties prenantes sont les particuliers, BCA ou les experts automobiles, une entité représentant la préfecture et les assureurs. Nous pouvons enregistrer les différents utilisateurs dans chaque organisation, ce qui nous donnera la possibilité après cette étape de définir les MSP.

##### Développement du Smart Contract

Les Smart Contract représentent les contrats virtuels ayant principalement un fonctionnement d’évènement déclencheur. Si les conditions d’une fonction sont validées alors il s’exécute.

Dans notre cas, un cas d'usage réalisable serait celui où un expert automobile certifie qu’un véhicule ne peux plus rouler. Alors il aurait la possibilité d’envoyer l’information « le véhicule ne peut plus rouler » qui sera immédiatement transmise chez les assureurs, ce qui annule les contrats d’assurances et facilite la transmission des informations offrant ainsi une réduction des coûts. Effectivement il y a un gain de temps et de ressources entre tous les différents acteurs de la chaîne et les étapes qui les séparent dans la compagnie, de même pour les transmissions de documents au gouvernement. Toutes les étapes avant le smart contract sont facilitées par la console IBM, contrairement aux autres blockchains qui doivent être codées manuellement du début à la fin. La console permet de générer toutes les étapes énoncées précédemment et de simplifier le processus.

##### Développement de l’application client

Une fois le smart contract réalisé, nous avons développé l’application « client » permettant de donner un aspect visuel au contrat réalisé en typescript.

Les deux différentes interfaces que nous avons créées sont liées aux différentes fonctionnalités. La première est celle de l’expert automobile, dans celle-ci il pourra remplir différents champs concernant les informations d’un véhicule à expertiser, la date, si la voiture est en état fonctionnel et s’il y a des réparations. La seconde est celle du ministère. Elle permet au gouvernement de certifier les différentes informations transmises par l’expert, pour que l’assureur soit ensuite informé des décisions prises (Voir Figures 3 et 4).

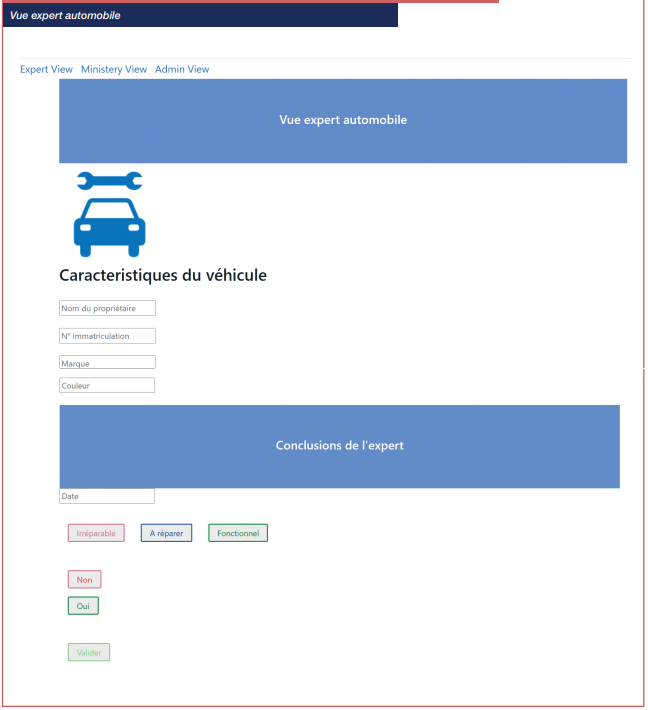


Figure 3 : Vue expert automobile

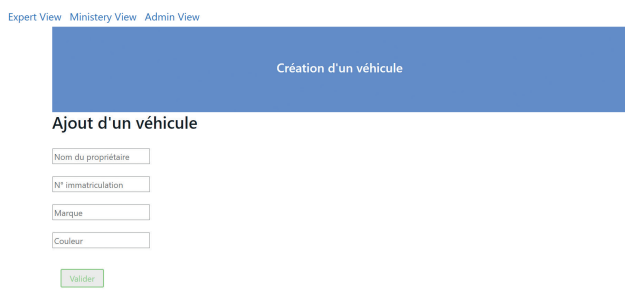


Figure 4 : Vue administrateur

Tout ce processus présente un gain de temps immense entre les différentes organisations ainsi qu’une sureté sans faille. Effectivement la sécurité des blockchains publiques ou privées permet d’affirmer qu’il est extrêmement difficile de falsifier les informations qu’elles contiennent. La cryptographie est un moyen difficile à inverser, ainsi avec le système d’imbrication entre les différents blocs, cela rend la corruption très complexe.

Avant de déployer l’application, la dernière étape est de relier le smart-contract et la console IBM à l’application client. Cette partie permet de récupérer les informations via les différentes « requêtes » faites depuis les vues clients. Les différentes couches sont dans l’ordre : l’application client, le backend puis la blockchain.

Finalement, nous pouvons voir l’ensemble du projet avec plusieurs entités : l’application client depuis laquelle l’utilisateur peut effectuer des requêtes qui seront envoyées à la deuxième entité, le « backend5 » ou autrement dit, le coté serveur.

Ce côté récupère les différentes requêtes et les traite selon les différentes fonctions développées. Le backend est en communication avec la base de données qui est développée avec MongoDB et avec la partie blockchain et ses smart-contracts. L’architecture globale de la blockchain est illustrée en Figure 5.

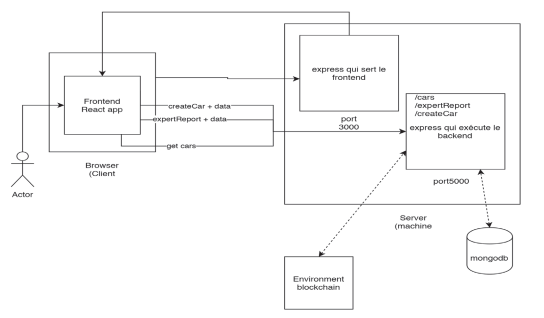


Figure 5 : Architecture globale de la blockchain

Après avoir développé l’intégralité de l’application client et ses différentes fonctionnalités, nous avons déployé l’application sur le « cloud » d’Amazon, cela permet aux consultants de l’entreprise de pouvoir accéder à cette démo sans notre présence et de s’en servir. Il s’agit d’Amazon web services (AWS) qui consiste en un ensemble de services proposés, comme des outils, de la puissance de calcul, des jeux ou encore du stockage. Le cloud est accessible par n’importe qui ayant une bonne connexion Internet.

#### Développement d’une architecture Micro-Services IA

L’IES, entité de la société BCA Expertise, opère plusieurs micro-services en interne basés sur l’Intelligence Artificielle dans le domaine de l’assurance automobile notamment, les services de sécurité, supervision, gestion des clients, stockage des données, Backup, service Mesh, développements IA ainsi que le développement des logiciels. Ces micro-services reposent sur des modèles entraînés au sein de l’IES. Ils sont accessibles via une API REST et sont exécutés sous forme de containers docker sur l’infrastructure de BCA Expertise.

Si aujourd’hui les solutions IA de l’IES rendent les services attendus, ils n’offrent pas les meilleures garanties en termes de scalabilité, de gestion multi-clients, de supervision et de bonnes pratiques de déploiement. De plus l’interaction entre les différents services n’est pas conçue de façon optimale. Souhaitant capitaliser sur ces micro-services, l’IES a décidé de les proposer en externe en mode SaaS.

Ainsi, l'objectif de ce projet est de concevoir et mettre en œuvre une nouvelle architecture, capable de supporter la charge induite par ce nouveau positionnement en tant que fournisseur de service et permettant d'intégrer rapidement de nouvelles offres de services, gérer l'accès et la facturation des clients ainsi que fournir tous les outils nécessaires au suivi de la production, tout en limitant au maximum l'impact de ce changement d'architecture sur les micro-services eux-mêmes.

Pour ce faire, il est nécessaire d’employer un socle technique commun permettant d’opérer des modules IA génériques mais également spécifiques à certains clients. L’unification des processus et l’industrialisation des déploiements permettent d’accepter plus facilement une montée en charge importante et de faciliter le développement de nouveaux services pour des clients de plus en plus nombreux.

Nous avons choisi d’employer le mode de fonctionnement Docker qui permet d’encapsuler un logiciel et toutes ses dépendances (bibliothèque, code source, fichiers de configuration) dans une entité appelée « container ». Ce container peut ensuite être déployé sur n’importe quel serveur disposant d’un environnement d’exécution compatible avec Docker.

Ce mode de fonctionnement facilite grandement le déploiement de logiciels. Avec Docker, le développeur livre un logiciel qui est à la fois « packagé » et prêt à l’emploi. Il ne reste plus à l’équipe d’exploitation de récupérer l’image du container du logiciel, de l’exécuter sur un ou plusieurs serveurs et de gérer les problématiques réseau pour permettre la communication avec d’autres systèmes ou l’accès réseau au logiciel déployé.

La Figure 6 ci-dessous présente l’intégration d’Istio dans un cluster Kubernetes pour implémenter l’architecture micro-services de la plateforme IA de l’IES. Une description détaillée de l’architecture est fournie en Annexe1.

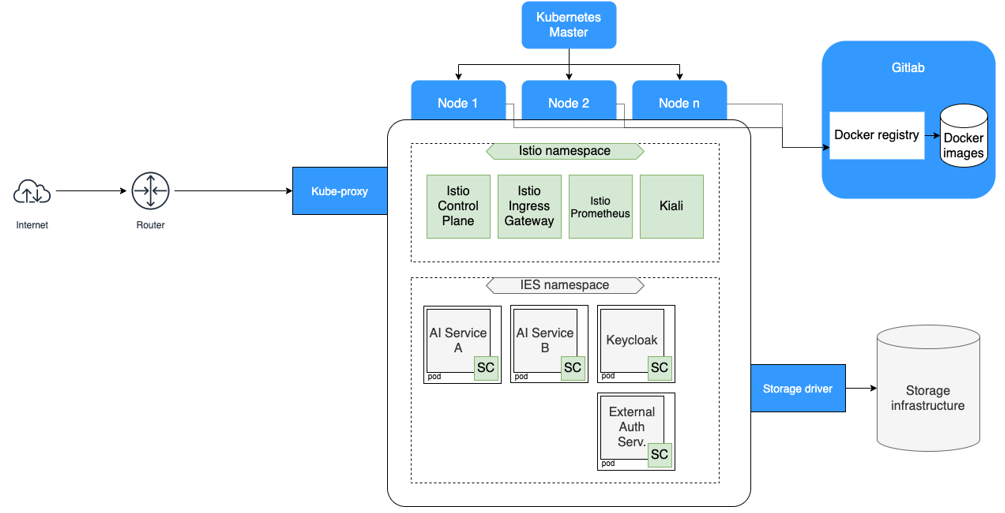


Figure 6: Implémentation de l’architecture micro-services de la plateforme IA de l’IES

L’architecture est conçue à partir de l’association de Kubernetes et Istio.

Kubernetes est une technologie très répandue dans les architectures microservices, en particulier dans les environnements cloud ou pour le déploiement de cloud native applications. Un cluster Kubernetes permet d’industrialiser le déploiement de containers Docker et de répondre à de nombreuses exigences d’un système en production :

* Scalabilité ;
* Optimisation des ressources ;
* Tolérance aux pannes ;
* Standardisation des déploiements (manifestes) ;
* Extensibilité (Custom Resource Definitions) ;
* Stabilité (rolling upgrades).

Istio est un Service Mesh utilisé depuis plusieurs années en production. Outre le fait de jouer le rôle d’API Gateway dans l’architecture microservices, Istio apporte des fonctionnalités qui viennent compléter celles de Kubernetes :

* Gestion et administration centralisée du maillage de services sans impact sur le développement des microservices ;
* Routage dynamique ;
* Contrôle d’accès ;
* Audit / monitoring / métriques / traçage
* Équilibrage de charge entre les multiples instances des microservices ;
* Circuit-breaking en cas d’indisponibilité d’un microservice pour éviter les erreurs à répétition ;
* Politiques de « retry » pour retenter l’appel à un service en cas d’erreur ou d’indisponibilité de celui-ci avant de considérer la requête en erreur.

#### Fragilité financière

Dans un contexte de crise sanitaire ayant déjà fragilisé la santé économique de beaucoup d’entreprises, il était important que les équipes de management du risque et du recouvrement redoublent d’efforts afin que leurs clients paient en temps et en heure et n’entraînent ainsi pas de problématiques de trésorerie supplémentaires. Bien prédire, ou mieux prédire, ses futurs mauvais payeurs devient par conséquent un enjeu incontournable.

L'opérateur Télécom La poste Mobile pour lequel l'algorithme a été réalisé commercialise des abonnements avec smartphone. Une part importante de ses clients est en fragilité financière et nécessite d'être relancée tous les mois en vue du paiement des factures. L'opérateur souhaitait mieux connaître le risque client au moment de la souscription afin de demander des garanties ou d’orienter les clients les plus à risque vers des offres plus adaptées.

Le projet vise à limiter les demandes de garanties aux seuls clients les plus à risque ainsi qu’à réduire le taux de résiliation des abonnements à 3 mois et atteindre un taux de 2% (le taux de résiliation actuel étant de 4%). Les gisements de données constituent désormais une opportunité permettant à la fois une connaissance approfondie des clients et prospects, mais en même temps une difficulté accrue pour les équipes d’analyser les risques et de proposer une méthode harmonisée de prévention et de gestion des impayés.

Les actions possibles sont nombreuses et complexes à coordonner pour être ajustées à chaque niveau de risque client : orientation sur des moyens de paiement moins risqués, recours à des solutions de scoring de fraude selon le canal de vente, optimisation de la date de prélèvement, pratique d’une avance sur consommation avant la contractualisation, limitations provisoires de service, rééchelonnement des montants dus, accompagnement par le conseiller clientèle vers des niveaux de services plus adaptés à la situation de l’abonné.

Afin de mener à bien cette mission, l’implémentation d’algorithmes de prédiction du niveau de risque pour un client à générer un impayé passe par les étapes suivantes.

##### Définition de la problématique et de l’indicateur métier

Contrairement à l’acte de fraude qui est parfois ambigu et difficile à déceler, l’impayé est un acte immuable. La qualité d’un client, pour lequel un historique à être « bon payeur » ou « mauvais payeur » est disponible, peut donc être évaluée assez rapidement. Ainsi, une approche supervisée est privilégiée plutôt qu’une méthode non supervisée souvent moins performante.

Afin de délimiter complétement la base de données d’apprentissage, il faut déterminer l’historique nécessaire de données étiquetées. Cette fenêtre de temps nécessaire dépend du volume de nouveaux abonnés mensuels de l’entreprise, et est fonction de la capacité à mettre à disposition ces données étiquetées. Il va de soi que l’apprentissage du modèle prédictif est d’autant meilleur que la base de cas de « bon payeur » et « mauvais payeur » est étayée. Il est également intéressant d’avoir un historique suffisamment long pouvant faire apparaître des effets de cycles ou saisonnalité. Pour ces raisons, nous avons donc choisi de nous baser sur un historique de deux ans de données clients pour notre approche de scoring avec l’opérateur téléphonique.

##### Pré-traitement et sélection des données pertinentes

Prédire le caractère « mauvais payeur » suppose d’avoir les données les plus pertinentes possibles. Voici quelques exemples de types de données qu’il est intéressant d’intégrer selon le cas pour avoir une vision complète des abonnés :

* Données sociodémographiques : âge, sexe, nationalité, langue, situation civile et familiale, CSP, lieu de vie, etc. ;
* Données comportementales liées aux goûts, choix et habitudes de consommation : historique des produits achetés ou services souscrits, options souscrites, fréquence d’achat et niveau d’utilisation des services, canal de souscription, etc. ;
* Données financières : banque, choix du mode de paiement, jour de prélèvement, etc. ;
* Données externes et enrichissement via des *open datas* :
  + Données macro-économiques territoriales :

L’INSEE publie tous les trois ou quatre ans un certain nombre d’indicateurs (éducation, chômage, revenus, démographie) à l’échelle de l’IRIS par exemple. Les *Îlots Regroupés pour l'Information Statistique* sont une maille élémentaire infra-communale de découpage du territoire français qui correspondent à des sortes de quartiers d’environ 2 000 habitants.

* + Données épidémiologiques:

La conjoncture sanitaire actuelle nous force à changer notre paradigme et notre façon de penser afin d’intégrer le risque de crise sanitaire dans la prise de décision. La souscription d’un service ou sa mise en œuvre peut être fortement impactée selon le département. Il est aussi nécessaire de prendre en compte leur impact pour actualiser le modèle, notamment celui sur les effets liés aux séries temporelles.

Une fois que toutes les variables jugées pertinentes par la prédiction ont été définies, nous utilisons des méthodes de réduction de dimensions ou de sélection de variables afin de ne sélectionner que les variables qui sont réellement les plus pertinentes statistiquement parlant.

Les méthodes de réduction dimensionnelle de type PCA (*Principal Component Analysis* ou ACP pour *Analyse en Composantes Principales* en français) sont très pratiques mais créent des combinaisons linéaires des variables existantes qu’il est parfois difficile d’interpréter. C’est pourquoi, pour une meilleure interprétabilité des résultats, nous avons privilégié l’utilisation d’une méthode de sélection de variables : tests statistiques (ANOVA, Chi2, t-test de Student…), suppression de variables colinéaires, *Forward selection*, *Backward elimination*, etc.

##### Apprentissage automatique

Lors de la phase du scoring du risque d’être « mauvais payeur », nous avons opté pour des méthodes ensemblistes et testé des algorithmes de *Random Forest* et *XGBoost*. C’est finalement le Random Forest qui nous a donné les meilleurs résultats.

#### Prédiction et pilotage des leads

L'objectif du projet est de développer une application pour notre Client SKODA permettant de prédire le nombre de leads qualifiés dans le secteur de l'automobile en fonction du plan marketing (publicité, messages, budget par partenaire), des éléments externes (ex. nombre de cas de COVID) ainsi que des caractéristiques de la gamme (modèles, âge, etc.,).

L'originalité du projet réside dans l'intégration de données externes, la prise en compte du plan marketing ainsi que la sélection des leads qualifiés à la place des leads d’une faible probabilité de transformation.

Sur le plan du pilotage, le caractère innovant réside dans le choix des KPI et l'ergonomie de la solution. Celle-ci a reçu le prix du Data Festival comme meilleur dispositif de pilotage.

Pour se faire nous avons commencé par la collection et la préparation des données. Tout d’abord, nous avons déterminé une date de début et de fin pour les campagnes fil rouge performance par la récupération des données sur les impressions. Nous avons récupéré par la suite des budgets par campagne fil rouge performance. Nous avons toujours des discussions en cours avec Remind sur les délais de mise à disposition de ces données.

Concernant l’algorithme d’apprentissage, nous avons choisi les modèles ensemblistes Random Forest et XGBoost vu leur bonne « explicabilité » et la précision des prédictions que ces algorithmes produit, associés au fait qu’ils permettent d’éviter le surapprentissage.

Afin d’évaluer nos modèles, nous avons comparé les résultats des algorithmes avec une valeur de référence (baseline) qui correspond à une prédiction évidente. Nous obtenons un coefficient R2 de 0,8, rendant les deux modèles satisfaisants. Le tableau ci-après détaille les résultats obtenus.

Une image contenant table

Description générée automatiquementFigure 5 : Résultats des algorithmes Random Forest et XGBoost

Baseline choisi : Le volume de leads qualifiés par partenaire et levier prédit correspond au volume de leads du mois M-1 pour le même levier et partenaire.

Moyenne mensuelle des leads : 100

### Conclusions

Soucieux de la qualité de nos solutions, nous restons à l’écoute de nos clients et faisons preuve de proactivité pour anticiper leurs besoins afin de proposer des expériences ergonomiques et des performances techniques toujours améliorées.

Les travaux menés en 2020 nous ont permis de développer quatre outils technologiques spécialisés pour nos clients dans les domaines de télécom et assurance automobile.

Des considérations d’exhaustivité des fonctionnalités, d’ergonomie d’usage et de performance ont guidé nos développements. Le choix de nouvelles technologies (langage de programmation, hébergement…) sert également notre objectif de différenciation par rapport à nos concurrents grâce à des plateformes hautement automatisées et robustes, capables d’offrir les meilleures performances tout en évoluant avec un volume de demandes en croissance.

# ANNEXES

Annexe 1 : Plateforme micro-services IA