

**100% ACTUAIRES &
100% DATA SCIENCE**

INSTITUT DES
ACTUAIRES

Voitures connectées : Comment améliorer la sinistralité ?



16 Novembre 2018
Hôtel Marriott Rive Gauche
Paris 14ème

//galea



- ∥ Introduction : Partenariat Galea/Ellis-Car
 - Présentation du cabinet Galea
 - Présentation de la Start-up Ellis-Car et de l'appli

- ∥ Premier axe : enrichissement des données
 - Démarche suivie
 - Enrichissement des données du portefeuille avec les données de la start-up
 - Résultats des algos avant et après enrichissement des données
 - Données fournies par Ellis-Car
 - Résultats et analyses sur l'apport des données externes présentés

- ∥ Deuxième axe : Prévention
Prévention et réduction de la sinistralité

- ∥ Conclusion et perspectives

Démarche de la tarification en assurance non-vie

Objectif

La tarification en assurance non-vie consiste en la modélisation de la fréquence de survenance des sinistres ainsi que son coût en fonction des différentes caractéristiques du contrat (profil de l'assuré, type de véhicule...)



Etude de la qualité du portefeuille



Détection des anomalies



Retraitement de la base de données



Retraitement des anomalies



Analyse univariée des variables tarifaires



Vérification de l'intérêt tarifaire des variables



Sélection des variables utiles à la tarification



Etude de corrélation et méthodes de sélection en régression



Modélisation de la prime pure



GLM:
Fréquence*Coût moyen



Préparation des données

- Compréhension des données
- *Feature engineering*
- Construction de la base finale
- Ecrêtement
- Traitements des données manquantes/incohérentes
- Statistiques descriptives
- Etude de corrélation



Modèles

- Base apprentissage / Base test
- Modèles appliqués (**GLM**, **CART**, **Random Forest**, **XGBoost**)
- Importance des variables
- Mesure de la pertinence des modèles (**RMSE**)
- Résultats / analyses



Intégration des données télématiques

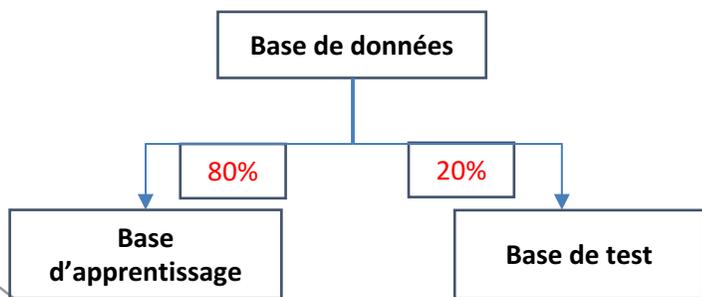
- Géocodage des codes postaux
- Traitement des données télématiques
- Rattachement à la base de données
- Impact sur les modèles



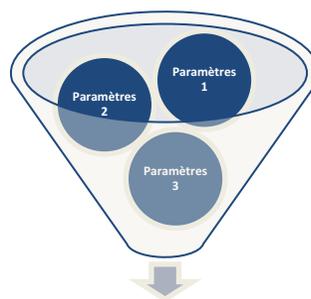
Données étudiées

- \\ Portefeuille automobile ;
- \\ Garantie de responsabilité civile ;
- \\ Données disponibles de qualité.

Démarche appliquée



1 Base d'apprentissage / Base test



2 Calibrage du modèle sur la base d'apprentissage



RMSE

3 Validation du modèle sur la base test

Algorithmes testés

- \\ Modèle linéaire généralisé (**GLM**).
- \\ Arbres de décision (**CART**).
- \\ Forêts aléatoires ou **Random Forest**.
- \\ eXtreme Gradient Boosting Machine (**XGBoost**).

Modèle	Vitesse d'apprentissage	Facilité d'explication de l'algorithme	Facilité de paramétrage	Pouvoir prédictif	Interprétabilité des résultats
GLM	++++	++++	++++	++	++++
CART	++++	++++	+++	++	+++
Random Forest	++	+	++	+++	+
XGBoost	+	+	+	++++	+

Qu'est ce que la télématique

- \\ La télématique de véhicule est la technologie d'enregistrement, d'envoi, de réception et de stockage des informations via des appareils de télécommunication.
- \\ Cette technologie permet de mieux évaluer les risques et assure un travail de prévention beaucoup plus poussé.
- \\ En assurance automobile, cette technologie permet de contrôler en temps réel le comportement du conducteur à travers des boîtiers par exemple.
- \\ Il existe plusieurs assureurs en France qui proposent des offres du type « pay as you drive »

Acquisition des données télématiques

- \\ Utilisation de nouvelles sources de données avec l'avènement du big data
- \\ Possibilité de qualifier plus finement les risques
- \\ Disponibilité de plusieurs sources de données : INSEE, OpenStreetMap, data.gouv...
- \\ Partenariat avec une start-up.



Dans le cadre de la présente étude, partenariat avec l'entreprise **Ellis Car**



CHALLENGE AXA BIG DATA ENTREPRISE INNOVANTE DES POLES PARIS REGION ENTREPRISES STASCOLE ECOLE POLYTECHNIQUE TELECOM PARIS TECH

A.I. & BIG DATA AUTOMOBILE

ARTIFICIAL INTELLIGENCE & AUTOMOTIVE BIG DATA

1,25 Million
Millions de morts dans le monde



WHO Statistics - 2016

Une Équation Simple



Comportement Humain

Sécurité Active

Infrastructure Routières

Une Équation Simple



Comportement Humain

cause
98%
des accidents

Nos solutions: Intelligence Artificielle & Big Data



Comme ça marche ?

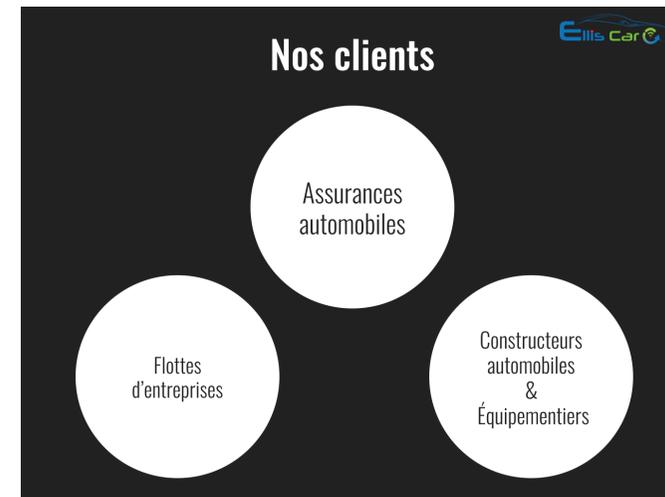


- 1 Smartphone
+ Beacon 
- 2 Devices Standalone
 
- 3 Connexion Native
  

The Ellis-Car Beacon



- The Beacon is a Bluetooth transmitter, equipped with an internal battery, it sends 24/7 a unique identifier.
- It is to be left in the glove compartment of the vehicle, and will help to record better quality data of each trip.
- Combined with a smartphone, this device allows to identify not only the driver, but also the car driven.
- It also allows to find where a car is parked, even when it was parked by another driver.
- In case the car is stolen, the beacon allows to find it back very quickly, thanks to an intelligent collaborative research system.





Reconnaissance Internationale

Ellis Car

-  **INNOVATION 2030**
CONCOURS MONDIAL D'INNOVATION
-  **CHALLENGE AXA BIG DATA**
-  **PARIS REGION ENTREPRISES**
-  **MAKE IT DRIVEABLE**
-  **PRIX INNOVATION 2017**
SÉCURITÉ ROUTIÈRE
TOUS RESPONSABLES

Avril 2014: Top 100 des start-ups les plus innovantes au monde

Novembre 2015: 2 Millions de constats d'accidents analysés

Avril 2016: Prix entreprises région Île de France

Septembre 2016: #1 Prix Ford parmi 300 entreprises

Mars 2017: Ministère de l'Intérieur: Sécurité Routière 2017

Exemple des variables télématiques récupérées :

- \\ Distance totale de routes en mètres ;
- \\ Distance d'autoroutes en mètres ;
- \\ Vitesse moyenne ;
- \\ Distances de routes recouvertes de pavés, de béton ou de bitume en mètres ;
- \\ Nombre d'accélération, de freinage brusques ;
- \\ Nombre d'excès de vitesse ;
- \\ Nombres de feux rouges, de stop, de passages piétons, de ralentisseurs... ;
- \\ Densité de routes, de la population vivant dans cette zone ;
- \\

Acquisition des données télématiques

Extraction et géocodage de la variable « Code Postal »

Code postal	Classe_SRA	Autres variables
AAAAA
BBBBB
CCCCC



Code postal	Latitude	Longitude
AAAAA	a	b
BBBBB	c	d
CCCCC	e	f

Transmission des codes postaux géocodés à Ellis Car afin de permettre l'extraction des données télématiques

Code postal	Latitude	Longitude	mean speed	stops	motorwayDistance	Autres variables
AAAAA	a	b	xxx	xxx	xxx	xxx
BBBBB	c	d	xxx	xxx	xxx	xxx
CCCCC	e	f	xxx	xxx	xxx	xxx

Jointure de la base télématique traitée à notre portefeuille RC auto en usage professionnel à l'aide de la clé « Code Postal ».

Code postal	Classe_SRA	Autres variables	mean speed	stops	motorwayDistance	Autres variables télématiques
AAAAA	xxx	xxx	xxx	...
BBBBB	xxx	xxx	xxx	...
CCCCC	xxx	xxx	xxx	...

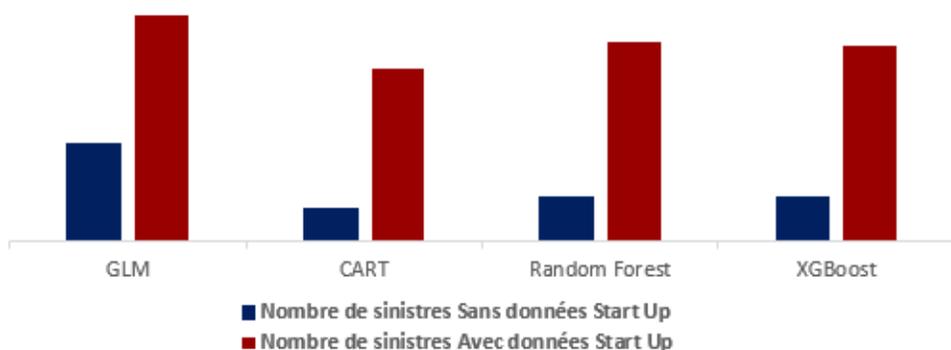
Traitement des codes postaux manquants :

- Création de la variable « Département » qui prend les 2 premiers chiffres du code postal;
- Affectation de la moyenne du département aux codes postaux manquants et ce, pour chaque variable télématique.
- En cas d'absence du département pour un certain code postal, lui affecter la moyenne de tous les codes postaux présents dans la base.

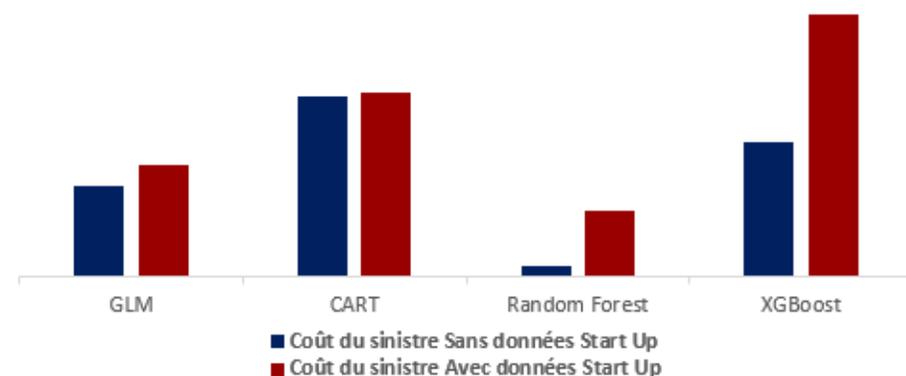
Résultats obtenus

- Une amélioration de la qualité de prédiction avec l'ajout des données télématiques.

Qualité de la prédiction (R^2)
sur le nombre de sinistres

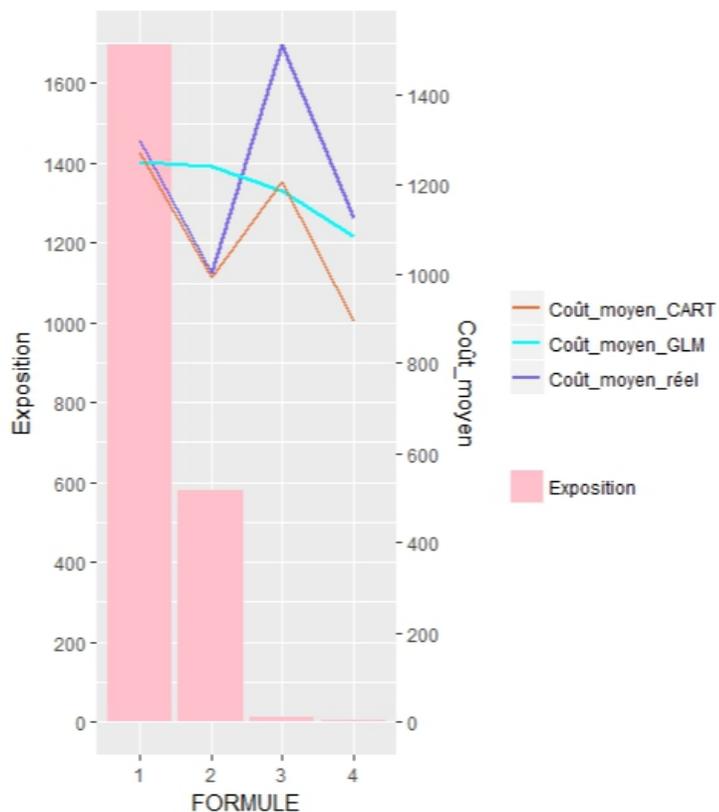


Qualité de la prédiction (R^2)
sur le coût du sinistre



Résultats obtenus

- Une analyse de l'importance des variables suivant les modèles afin d'affiner les prédictions.



LA PREVENTION A L'AIDE DES TELEMATIQUES

- Big data et analyses prédictives
- Analyse du comportement des conducteurs
- Réduction du risque d'accidents en évitant les accidents de la route
- Ellis Car primé en 2017 par le ministère de l'intérieur en matière de sécurité routière
- ...

CONCLUSION

L'adoption des technologies télématiques a des avantages à différents niveaux :

\\ National

- Baisser la fréquence des accidents,
- Localiser les véhicules volés,
- Réduire le temps de réponse lors de la survenance d'un accident,
- ...

\\ Assureurs

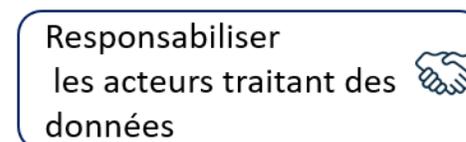
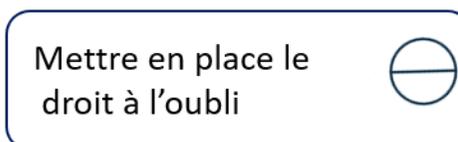
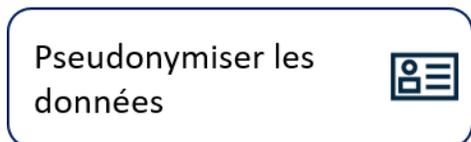
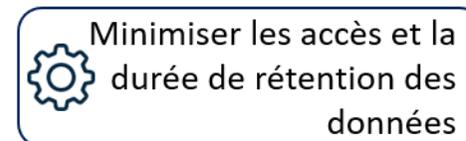
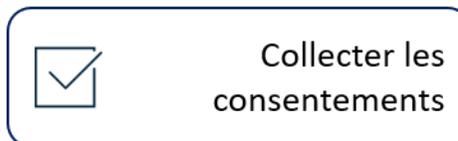
- Renforcer l'exactitude des primes,
- Réduire les coûts des sinistres,
- Meilleure connaissance du risque,
- ...

\\ Assurés

- Réduire les primes,
- Adopter un comportement plus prudent,
- Profiter d'autres services comme les services d'urgence, de géolocalisation, conseil de conduite,
- ...

CONCLUSION

- \\ Nécessité de connaître les limites de la collecte, la légitimité de l'accès et le traitement des données.
- \\ La protection des données à caractère personnel est largement affirmée par le droit :
 - La CNIL : « Pack de conformité : Véhicules connectés et données personnelles »
 - RGPD : Impose de nouvelles règles de sécurisation et d'utilisation des données



CONCLUSION

- \\ Une amélioration du tarif à mettre en regard au coût d'intégration des données ;
- \\ Volonté des assureurs à utiliser des algorithmes « boîte noire » ;
- \\ Application sur l'ensemble des garanties (dommages, vol, ...) ;
- \\ Analyse plus fine de la rentabilité et amélioration de la valeur client ;
- \\ ...

MERCI POUR VOTRE ATTENTION !

QUESTIONS

