



Tarification des contrats automobiles à l'aide du Machine Learning

Présentation de Mazars Actuariat :

Mazars Actuariat est une entité du groupe Mazars, composée d'experts techniques de haut niveau et spécialisée dans les missions d'Audit et de Conseil dans le secteur de l'Assurance et de la Banque. Elle regroupe les expertises suivantes : Actuariat d'Assurance, Ingénierie Financière et Engagements Sociaux. Ce projet s'effectuera dans le pôle Actuariat d'Assurance, qui intervient sur des missions variées : conception et validation de modèles prospectifs (EV, BE, SCR), évaluation et revue de provisions techniques, mise en place des normes IFRS, tarification, statistiques, etc. Il dispose d'un large éventail de clients parmi les plus gros assureurs français, à la fois en assurance-vie et non-vie.

Contexte et Enjeu :

En assurance IARD (Incendies, accidents, risques divers), la tarification des contrats joue un rôle incontournable du fait qu'elle assure un équilibre entre les primes à percevoir et les sinistres à indemniser.

Dans le contexte assurantiel de l'automobile actuellement très concurrentiel, les acteurs tentent de se démarquer en faisant appel à des méthodes de plus en plus novatrices. En effet, à l'affût des garanties les plus avantageuses, les assurés n'hésitent pas à changer de compagnie quand leurs primes ne reflètent pas leur perception de leur conduite ou si elles sont trop coûteuses. Il en va de soi qu'il faut ainsi segmenter autant que possible les grilles tarifaires pour capter le coût de couverture de chaque assuré. C'est dans cette lancée que nous allons utiliser les modèles de Machine Learning pour tarifier des contrats d'assurance automobile.

Problématique et objectif :

Nous disposerons d'une base étendue regroupant les informations de milliers d'individus qui vous sera ultérieurement transmise. Il s'agira dans un premier temps d'en effectuer une analyse descriptive pour connaître les profils majoritaires contenus dans le portefeuille d'étude.

Ensuite, nous utiliserons la méthode « Fréquence x Coût Moyen » qui consiste à modéliser d'une part la fréquence de sinistres et d'autre part le coût des sinistres. Cette démarche pourra se faire en deux sous-parties. On pourra utiliser pour chaque facteur les modèles linéaires généralisés, les modèles d'apprentissage supervisés à base d'arbres (CART, Bagging, Boosting, RandomForest) ou autres types d'algorithmes jugés pertinents.

Enfin, nous choisirons des critères de sélection de modèles pour chacun des facteurs avant de procéder au produit des deux modèles qui se seront avérés les plus pertinents.

Eléments de bibliographie :

- Michel Denuit, Arthur Charpentier, *Mathématiques de l'assurance non-vie : Tome 2, Tarification et provisionnement*
- Marie-Julienne Diouf, *Modélisation de la résiliation pour non-paiement de primes en assurance automobile*, mémoire présenté en 2019 devant l'Institut des Actuaire

Contacts :

Marie-Julienne DIOUF : marie-julienne.diouf@mazars.fr

Imen SAID : imen.said@mazars.fr