

**PROJET D'INTELLIGENCE ARTIFICIELLE APPLIQUEE A L'ACTUARIAT - IMAFA 2020-2021**  
**PREDICTION DE LA PRIME A PARTIR DES OBSERVATIONS DU MARCHE**

**Marius AKRE**

Entrepreneur en Assurance, vous désirez commercialiser certains produits d'assurance. Cependant vous n'avez pas encore le budget nécessaire pour employer un spécialiste de calcul des primes avec les formules actuarielles. Cependant, vous pouvez aisément et à moindre coût observer les prix des concurrents pour ces produits sur le marché. Vous décidez alors d'utiliser un logiciel d'Intelligence Artificiel (IA) pour qu'il puisse apprendre le modèle de prix sous-jacent, afin que vous ayez un outil qui vous permettra de fournir une prédiction fiable de la prime à nouveau client.

Dans le cadre de ce projet, le scope sera restreint au produit d'assurance suivant : **Temporaire décès**.

- Dans un premier temps, il vous faudra générer le dataset dont vous aurez besoin pour l'IA. Ce dataset sera généré à partir des formules actuarielles de la prime annuelle d'assurance de la temporaire décès pour un set limité de paramètres. Cela correspondra aux prix observés sur le marché. Les tables de mortalité à utiliser sont les dernières tables en vigueur.

$${}_nA_x = \sum_{k=0}^{n-1} v^{k+1} * {}_{k|1}q_x$$

|                |                                      |
|----------------|--------------------------------------|
| $m (m \leq n)$ | $P = \frac{{}_nA_x}{\ddot{a}_{x:m}}$ |
|----------------|--------------------------------------|

Le set de paramètres à utiliser est l'ensemble des combinaisons possibles pour les valeurs suivantes :

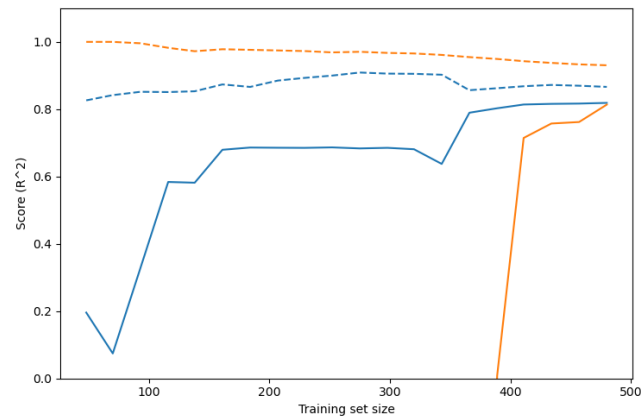
| Age | Nombre de paiements de primes annuelles | Maturité | Taux d'intérêt | Montant garanti |
|-----|---|----------|----------------|-----------------|
| 20  | 1                                       | 1        | 0              | 500             |
| 30  | 5                                       | 5        | 0.5%           | 1,000           |
| 40  | 10                                      | 10       | 1.0%           | 2,000           |
| 50  | 20                                      | 20       | 1.5%           | 4,000           |
| 60  | 30                                      | 30       | 2.0%           | 8,000           |
|     | 40                                      | 40       | 2.5%           |                 |

Nous aurons donc un dataset avec :

- 5 caractéristiques ("features") : âge, nombre de paiements de primes annuelles, maturité, taux d'intérêt, montant garanti,
- un nombre d'échantillons strictement (car le nombre de paiements de primes annuelles est inférieur ou égale à la maturité) inférieur à 5400 ( $5 * 6 * 6 * 6 * 5$ ).
- une cible ("target") : la prime annuelle.

- Ensuite, effectuer une **analyse des données** obtenues. Par exemple : count, moyenne, médiane, quantiles, histogramme des caractéristiques ; évolution de la prime en fonction de chaque paramètre, évolution de la prime en fonction de couple de paramètres, graphiques d'analyse ; vérification d'absence de données aberrantes, manquantes, etc.
- Séparez vos données en set de **test, de validation et de tests**.

- Choisissez et entraînez vos modèles en argumentant/explicant la démarche suivie/vos choix.
- Représentez les **courbes d'apprentissage** de vos différents modèles.



- Donnez la **conclusion** de votre étude par rapport à l'objectif de départ : avoir un outil d'IA qui vous permet de fournir une prédiction fiable de la prime pour un nouveau client, à partir des observations sur le marché et, sans avoir le modèle actuariel de tarification sous-jacent.
- Votre logiciel devra permettre à l'utilisateur d'utiliser une **interface pour avoir une estimation de la prime** en fonction des paramètres saisis. Le test score du modèle d'IA devra aussi être affiché, afin de pouvoir avoir une idée de la fiabilité de la prévision.

## Bonus

Certains éléments, s'ils sont pris en compte dans votre projet, seront sujets à bonus :

- **Digitalisation** : vos solutions logicielles sont utilisables directement sur une page Web (via un PC quelconque ou un mobile) pour faire une prédiction.

## REFERENCES

- Théorie et pratique de l'assurance vie, 3ème édition. Pierre Petauton
- Python for Data Analysis. Wes McKinney.
- Python Data Science Handbook. Jake VanderPlas
- Le Machine Learning avec Python. Andreas C. Muller et Sarah Guido



## DEADLINES

- Vendredi **4 décembre** : chaque groupe passera présenter sa version **DRAFT 1** (5 minutes),
- Vendredi **11 décembre** : chaque groupe passera présenter sa version **DRAFT 2** (5 minutes),
- Mercredi **16 décembre** : envoi de la **version finale** par mail (**logiciel/script + rapport**).