

Quadrature du Net

A l'attention de Madame Noémie LEVAIN

Madame, Monsieur,

La Cnaf a pris connaissance de l'avis de la Commission d'accès aux documents administratifs n°20226179 du 15 décembre 2022 avec intérêt.

Nous relevons que la Commission a émis un avis favorable à la communication des documents suivants demandés par la Quadrature du Net :

1. Communication, pour les anciennes versions de l'algorithme, du détail de la formule du calcul du score de risque utilisé par la Cnaf dans le cadre de sa politique de contrôle ;
2. Communication des documents échangés avec la CNIL, qui ne rentrent pas dans le cadre des formalités préalables aux obligations prévues par la réglementation relative à la protection des données (RGPD et Loi Informatique et Libertés).

Par ailleurs nous constatons que la CADA a émis un avis défavorable à la communication du modèle en vigueur, construit en 2018.

Avant ce modèle, seuls deux modèles ont été utilisés pour calculer des niveaux de risque pour l'ensemble des allocataires des Caf : celui construit en 2010, remplacé ensuite par celui construit en 2014. Nous vous communiquons pour chacun, algorithme et tableau de variables.

Les algorithmes retracent toutes les opérations menées pour aboutir à la sélection des variables des modèles, y compris la construction des données d'apprentissage, non retenues lors des modélisations statistiques.

A noter qu'en 2010 l'algorithme produisait cinq calculs de niveaux de risque, le niveau le plus élevé était alors retenu. Le modèle de 2014 ne produisait qu'un seul niveau de risque.

Les tableaux précisent les variables utilisées dans les calculs de niveaux de risque, ainsi que leurs modalités et les odd-ratios associés.

Les modèles utilisés en 2010, puis en 2014, exploitaient quelques variables qui sont toujours présentes dans le modèle en vigueur. Par extension de la décision de la CADA concernant le modèle en vigueur, nous ne pouvons pas vous communiquer ces données sans porter atteinte



à la politique de lutte contre la fraude sociale. Ainsi, les algorithmes et tableaux ont été purgés de 6 variables concernant le modèle 2010 et de 3 variables pour le modèle 2014.

S'agissant des documents échangés avec la CNIL et ne rentrant pas dans le cadre des formalités préalables aux obligations prévues par la réglementation relative à la protection des données (RGPD et Loi Informatique et Libertés), nous ne disposons pas de tels documents et ne serons donc pas en situation de répondre à votre demande.

Cette impossibilité s'explique notamment par le délai écoulé depuis la mise en œuvre du traitement et les changements d'organisation au sein de la Cnaf depuis cette date.

Vous trouverez toutefois en pièce jointe :

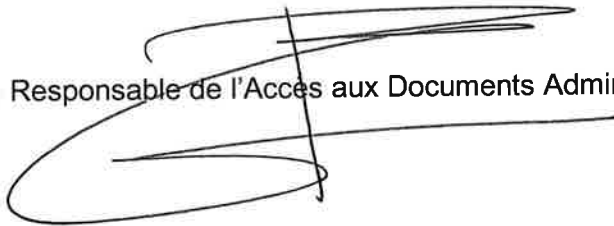
- Les algorithmes des deux anciens modèles datamining de 2010 et 2014
- Les tableaux de variables de ces deux modèles enrichis des odd ratios

Les agents de la Cnaf en charge de ce dossier, restent disponibles pour vous apporter des réponses complémentaires.

Veillez agréer, Madame, Monsieur, mes salutations distinguées,

Paule-Marie Grégoire,

Personne Responsable de l'Accès aux Documents Administratifs

A handwritten signature in dark ink, consisting of a large, stylized 'P' followed by a horizontal line and a vertical stroke.