

Nous réalisons un exercice pour vérifier la puissance des IA à résoudre des questions écologiques à partir d'un jeu de données. En ne considérant pas les étapes déjà réalisées précédemment et en se donnant carte blanche sur les méthodes et codes à produire, je te demande de répondre à l'exercice décrit ci-après en écrivant un code R permettant de répondre aux 4 missions à la fin de l'énoncé.

Début de l'énoncé de l'exercice : « Pour la partie analyse de données comme telle, vous travaillerez à prédire la santé des individus d'une espèce d'oiseaux omnivores de taille moyenne, capturés dans des grands parcs urbains, à partir d'une série de variables décrivant leur environnement.

La santé de chaque individu (`bird_health_index`) est mesurée à l'aide d'un indice de masse corporelle. Les valeurs positives présentent des oiseaux plus gras/en santé que la moyenne, les valeurs négatives des oiseaux plus maigres, moins en santé.

Pour tenter d'expliquer la santé des individus, les variables suivantes ont été mesurées, centrées sur chacun des oiseaux, dans un rayon de 500 m de la capture :

Le nombre de mangeoires (`feeder_count`)

La densité de routes (`road_density`)

Le nombre de poubelles (`thrash_can_count`)

La densité d'arbustes (`shrub_density`)

La présence d'oiseaux de proie (`raptor_presence`)

Vous remarquerez que dans beaucoup de cas, certaines variables n'ont pu être mesurées pour des raisons d'accès, de sécurité, etc (NA). Pour chaque individu, vous avez aussi la ville (`city_id`) et le parc (`park_id`) où il a été capturé.

Vos missions sont donc les suivantes :

1-Explorer le jeu de données pour bien le maîtriser

2-Modéliser à partir de la base de données `dataset.csv`, avec la technique de votre choix, la santé de ces oiseaux représentée par la variable : `bird_health_index`

3-Déterminer, basé sur votre analyse, quels sont les prédicteurs les plus importants, parmi les 6 mesurés, pour comprendre la santé de ces oiseaux

4-Prédire à partir de la base de données `predict.csv` la santé des oiseaux des 10 individus pour lesquels vous n'avez pas de mesure de santé. »