

Projet GenIALearn : Intérêts et limites de l'apprentissage statistique et de l'apprentissage profond appliqués à la sélection génomique multi-caractères

Coordonné par Eric Barrey et Didier Boichard de l'unité GABI UMR1313

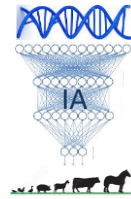
Financé par le métaprogramme INRAE DIGIT-BIO

L'unité GABI UMR1313 (INRAE, Domaine de Vilvert, Jouy-en-Josas), sous tutelle d'INRAE et AgroParisTech mène des recherches qui ont trait à la génétique des animaux d'élevage pour relever les défis sociétaux et environnementaux actuels. Notamment, GABI mène des travaux de recherche sur la sélection génétique de caractères phénotypiques d'intérêt économique chez les animaux d'élevage (résistance aux maladies, efficacité alimentaire, émission en méthane, qualité de la viande et du lait, performances sportives chez les chevaux, ...). Dans ce cadre, le projet GenIALearn explore les possibilités offertes par l'intelligence artificielle et notamment l'utilisation de modèles de deep learning et de machine learning pour prédire les caractères d'intérêt. Pour mener à bien ces travaux, il convient de disposer de très nombreuses données de génotypage (informations sur le génome) et de phénotypage (mesures des caractères d'intérêt) correspondant afin de réaliser correctement l'apprentissage profond des modèles. Dans les phase de mise au point des premiers modèles, nous utilisons des données humaines, pseudonymisées et randomisées à la source par UK Biobank, de génotypes et quelques caractères simples (poids, taille, mensurations,...). Ce n'est pas le caractère humain qui est utile dans ce travail de recherche méthodologique mais bien le plus grand nombre d'observations disponibles qui permet un meilleur apprentissage des modèles. Les données humaines utilisées dans le projet seront des données pseudonymisées déjà collectées et mises à disposition par la banque de données UK Biobank sous le projet approuvé N°96326. L'avantage de ces données est qu'elles sont nombreuses et bien gérées par UK Biobank notamment vis-à-vis des consentements éclairés des personnes volontaires et des règles de protection des données à caractère personnel en vigueur en Grande Bretagne, le lieu de collecte des données.

A INRAE, le projet de recherche est mené par un groupe d'une dizaine de personnes de deux unités INRAE (GABI UMR-1313, MIA-Paris-Saclay) et une unité universitaire (IBISC, Université Evry - Paris-Saclay) et coordonné par Eric Barrey (DR) et Didier Boichard (DR). La base légale de ce projet est l'exécution de notre mission d'intérêt public de production de connaissances.

Les données importées de UK Biobank seront conservées, dans les meilleures conditions de sécurité et de confidentialité, pendant toute la durée du projet de recherche à savoir 36 mois jusqu'au 23 Mars 2026 sauf si une prolongation est demandée à UK biobank afin de terminer le projet. Aucun transfert de données hors de l'Union européenne n'est réalisé. A l'issue de cette durée, les données importées de UK biobank seront détruites après la publication des travaux scientifiques.

Conformément au Règlement européen relatif à la protection des données personnelles et à la loi Informatique et Libertés, vous bénéficiez d'un droit d'accès, de rectification, d'effacement des



informations qui vous concernent. Si vous souhaitez exercer ces droits et/ou obtenir communication des informations vous concernant, vous pouvez vous adresser :

Préférentiellement à :

UK Biobank Limited
Units 1 & 2 Spectrum Way
Adswood, Stockport, Cheshire, SK3 0SA, UK
www.ukbiobank.ac.uk

ou à l'unité GABI en écrivant à :

INRAE
GABI UMR-1313
Bat 211
Domaine de Vilvert
78350 Jouy-en-Josas
jj-gabi-direction@inrae.fr

En cas de difficultés, vous pouvez également contacter le Délégué à la protection des données personnelles (DPO) accompagnant l'unité. Ses coordonnées sont :
[DPO INRAE](#), 24, Chemin de Borde Rouge – Auzeville- CS 52627 ; 31326 Castanet Tolosan Cedex ; France
Tél. : +33 1 (0)5 61 28 54 37 ; Courriel : cil-dpo@inrae.fr

Si vous estimez, après nous avoir contactés, que vos droits Informatique et Libertés ne sont pas respectés, vous avez la possibilité d'introduire une réclamation auprès de la CNIL par courrier postal : Commission Nationale de l'Informatique et des Libertés 3 Place de Fontenoy – TSA 80715 – 75334 PARIS CEDEX 07 ou en ligne <http://www.cnil.fr/>