

Agro Data Scientist : De la Biostatistique au Deep Learning pour le Vivant

Lien : https://afrique-formation.com/formation/agro-data-scientist-de-la-biostatistique-au-deep-learning-pour-le-vivant

DURÉE
10 jours (70h)

RÉFÉRENCE
AA05

CATÉGORIE
Agriculture Intelligente et Agriecologie

Afrique Formation

OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Maitriser les langages R et Python pour le traitement de donnees biologiques massives
- Concevoir des plans d'experiences robustes et en analyser la significativite statistique
- Developper des modeles predictifs (Machine Learning) adaptes a la variabilite du vivant
- Traiter des donnees spectrales (NIRS) pour le controle qualite non destructif
- Analyser et cartographier la variabilite spatiale des sols et des rendements (Geostats)
- Utiliser le Deep Learning pour automatiser le phenotypage par imagerie
- Creer des tableaux de bord interactifs pour l'aide a la decision agronomique
- Integrer des donnees heterogenes (Climat, Sol, Plante, Genotype) dans une approche systemique
- Evaluer la performance et la robustesse des modeles avant leur deploiement terrain
- Communiquer efficacement les resultats d'analyses complexes a des parties prenantes metier

POUR QUI ?

- Ingenieurs agronomes souhaitant acquerir une double competence en Data Science
- Data Scientists generalistes desirant se specialiser dans les sciences du vivant
- Chercheurs et doctorants en agronomie, ecologie ou genetique vegetale
- Responsables R&D dans l'industrie semenciere ou phytosanitaire
- Analystes de donnees au sein d'instituts techniques ou de chambres d'agriculture
- Responsables Qualite et Laboratoire en industrie agroalimentaire (volet Chimiometrie)
- Bio-informaticiens cherchant a elargir leur champ d'action vers l'echelle parcelle

Afrique Formation

AFRIQUE FORMATION
La Référence en Afrique pour la Formation



Programme détaillé

1 / Fondamentaux de la gestion de donnees agronomiques

- Structuration et nettoyage de bases de donnees complexes (Data Wrangling avec Pandas/Tidyverse)
- Gestion des valeurs aberrantes et donnees manquantes dans les essais biologiques
- Interoperabilite des formats (CSV, JSON, NetCDF) et bases de donnees SQL pour l'agronomie

2 / Biostatistiques et inference classique

- Conception de plans d'experiences (Blocs randomises, Split-plot, Carrés latins)
- Analyse de variance (ANOVA) et tests post-hoc pour la comparaison de varietes
- Regressions lineaires simples et multiples pour modeliser les interactions GxE (Genotype x Environnement)

3 / Analyse de donnees multidimensionnelles (Exploratoire)

- Analyse en Composantes Principales (ACP) pour la caracterisation phenotypique
- Analyse Factorielle des Correspondances (AFC) pour les donnees qualitatives
- Classification hierarchique (CAH) et K-means pour le zonage ou le typage varietal

4 / Introduction au Machine Learning pour le vivant (Supervisé)

- Algorithmes d'arbres de decision et Random Forests pour la prediction de rendements
- Support Vector Machines (SVM) pour la classification binaire (Malade/Sain)
- Evaluation des modeles : Matrice de confusion, ROC-AUC et validation croisee (Cross-validation)

Afrique Formation

5 / Machine Learning pour le vivant (Non-supervisé et Avancé)

- Gradient Boosting (XGBoost, LightGBM) appliqué aux donnees agro-climatiques
- Selection de variables (Feature importance) pour identifier les facteurs limitants
- Clustering DBSCAN pour la detection d'anomalies spatiales dans les parcelles

6 / Series temporelles et modelisation climatique

- Analyse de tendances et saisonnalite des donnees meteorologiques
- Modeles ARIMA et LSTM (Long Short-Term Memory) pour la prevision de recolte
- Calcul d'indices bioclimatiques cumules (GDD, stress thermique)

7 / Chimiometrie et analyse spectrale (Agroalimentaire)

- Pre-traitement des spectres proche infrarouge (NIRS) : Derivation, SNV, MSC
- Regression PLS (Partial Least Squares) pour la prediction de la qualite (Proteines, Brix)
- Discrimination de l'origine geographique des produits par analyse discriminante (PLS-DA)

8 / Geostatistiques et analyse spatiale

- Analyse de la structure spatiale par variographie experimentale
- Interpolation par Krigeage pour la cartographie de la fertilite des sols
- Gestion de l'autocorrelation spatiale dans les modeles agronomiques

9 / Deep Learning et Vision par Ordinateur (Computer Vision)

- Architectures de reseaux de neurones convolutifs (CNN) pour l'analyse d'images foliaires
- Detection d'objets (YOLO) pour le comptage de fruits ou d'adventices
- Segmentation semantique pour l'estimation de la surface foliaire (LAI)

10 / Analyse de donnees genomiques et selection assistee

- Introduction aux donnees omiques et matrices de genotypes (GWAS)
- Prediction genomique (Genomic Selection) via modeles lineaires regularises (Ridge, Lasso)

Afrique Formation

- Association genotype-phenotype pour l'identification de marqueurs d'interet

11 / Deploiement de modeles et outils d'aide a la decision (OAD)

- Creation d'applications web interactives avec R Shiny ou Python Streamlit
- Conteneurisation des modeles avec Docker pour la mise en production
- APIfication des algorithmes pour l'integration dans des outils tiers

12 / Analyse de survie et epidemiologie vegetale

- Courbes de Kaplan-Meier pour l'etude de la resistance aux maladies
- Modeles de Cox pour analyser les facteurs de risque phytosanitaires
- Simulation de la dispersion des pathogenes dans le paysage

13 / Ethique de la donnee et interpretation causale

- Distinction entre correlation et causalite dans les systemes biologiques complexes
- Biais algorithmiques et equite dans les modeles de recommendation agricole
- Souverainete des donnees agricoles et aspects reglementaires (GDPR)

14 / Projet integrateur : Du Dataset a la Valorisation

- Hackathon sur un jeu de donnees reel (ex: prediction qualite ble ou detection maladies)
- Pipeline complet : Nettoyage -> Modelisation -> Validation -> Visualisation
- Restitution et Storytelling de la donnee pour des decideurs non-experts

Afrique Formation

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Mises en Situation pour faciliter l'assimilation
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Prochaines dates programmées

15 au 26 Fév. 2027	Présentiel - Casablanca
12 au 23 Avr. 2027	Présentiel - Casablanca
07 au 18 Juin 2027	Présentiel - Casablanca
02 au 13 Août 2027	Présentiel - Casablanca
27 Sep. au 08 Oct. 2027	Présentiel - Casablanca

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Réservation & Renseignements

Téléphone : +212 522 247 210
Email : contact@afrique-formation.com
Web : https//www.afrique-formation.com

Afrique Formation — Tous droits réservés