

MACHINE LEARNING MODÈLE PRÉDICTIF AVEC LA RÉGRESSION LOGISTIQUE

Durée & Format

Durée : 14 heures

Format disponible : Présentiel ou Distanciel

Tarifs

Tarif public intra : Consulter notre [site Internet](#)

Tarif public inter : Consulter notre [site Internet](#)

Public visé

Toutes personnes souhaitant maîtriser l'utilisation d'un modèle prédictif basé sur la méthode de la régression logistique.

Prérequis

- Une connaissance des outils statistiques de base est souhaitée
- Une connaissance des approches prédictives classiques (régression) est un plus mais pas strictement nécessaire.

Objectifs pédagogiques

A l'issue de cette formation, l'apprenant sera capable de :

- D'identifier le contexte général d'utilisation de la régression logistique
- De connaître les concepts mathématiques inhérents à la régression logistique
- De mettre en œuvre et analyser les résultats (tableaux, graphiques) d'une modélisation de type régression logistique
- De calculer les probabilités ajustées d'apparition d'un événement « succès »
- De comparer la régression logistique avec d'autres outils type Afd, Méthode de classement
- D'interpréter les coefficients tels que les odds-ratios, notamment dans un contexte épidémiologique
- D'identifier et de résoudre les problèmes rencontrés lors de la mise en œuvre d'un modèle de type régression logistique
- De détecter et traiter les colinéarités éventuelles entre variables explicatives

Programme et déroulement

Jour 1 – Matin

Utiliser le modèle de régression logistique

- Variable explicative et variable expliquée (continue / binaire)
- Différences entre la régression linéaire classique et la régression logistique
- Variables explicatives qualitatives, variables explicatives quantitatives
- Objectifs de la régression logistique
- Définition du modèle Logit (courbe sigmoïde)
- Découpage en classes des variables explicatives quantitatives
- Conditions d'utilisation à respecter

Maîtriser les fondements mathématiques

ARKESYS.NET

70 rue Bergson – 42000 Saint-Étienne
04 28 95 15 82 - info@arkesys.fr
SAS AU CAPITAL DE 60 000 € - SIRET : 501 033 609 00030
ETS secondaire : 58 avenue Debouze - 69007 Lyon - SIRET : 501 033 609 00048
ETS secondaire : 293 route de la Seyne - 83190 Ollioules - SIRET : 501 033 609 00055



- Le modèle linéaire généralisé
- Le maximum de vraisemblance
- Recherche des coefficients du prédicteur linéaire
- Transformation des coefficients

Jour 1 – Après-midi

Mettre en œuvre et analyser des résultats d'un modèle de régression logistique

- Estimation et interprétation des coefficients du modèle
- Test de significativité du modèle (validation du modèle)
- Tests d'apport d'une variable (test de Wald, tests sur les rapports de vraisemblance)
- Interprétation du χ^2 de Wald
- Odds-ratios
- Parallèle odds ratios et risques relatifs
- Analyse du tableau de classement
 - Taux de réussite, taux d'échec
 - Vrais positifs, vrais négatifs, faux positifs, faux négatifs
- Probabilités ajustées et utilisation du modèle à des fins de prédiction
- Mise en œuvre et interprétation des résultats (tableaux, graphiques) de la régression logistique
 - Sur un tableau de contingence
 - Sur un tableau composé de variables explicatives uniquement qualitatives, uniquement quantitatives, qualitatives et quantitatives
- Modèle de régression logistique multinomial
- Mise en œuvre et analyse des résultats d'un modèle de régression logistique multiple
- Estimation et interprétation des coefficients du modèle multiple

Jour 2 – Matin

Mesurer la qualité d'un modèle de régression logistique

- Qualité d'ajustement du modèle (coefficients R^2)
- Qualité d'estimation des coefficients du modèle, qualité de prédiction
- Intervalles de confiance des coefficients du modèle
- Intervalles de confiance des odds-ratio
- Lien entre la qualité du modèle et :
 - Les effectifs des classes
 - La colinéarité des variables explicatives
- Sélection du modèle final
- Matrice de confusions

Analyser les variantes de la régression logistique

- La régression logistique multimodale
- La régression logistique ordinale

Jour 2 – Après-midi

Comprendre le sous et surajustement

- Notions de biais d'un modèle
- Notions de variance d'un modèle
- Optimalité variance & biais
- Jugement de la qualité d'un modèle
- Méthode du Data Train / Data test
- Validation croisée
- Classique LOO (leave One Out)

- Validation croisée par k fold

Réaliser des exercices pratiques

- Applications avec R
- Démonstration avec Excel

Cas d'études sur données des apprenants

Jeux de données

Afin de s'approcher au mieux des réalités quotidiennes des praticiens, nous suggérons de nous appuyer pour l'animation pratique de thématiques et surtout de jeux de données reflétant le quotidien des apprenants.

Cet élément est un facteur de réussite pour la formation. Elle permet aux apprenants de :

- Se "reconnaître" dans les thèmes abordés,
- Mieux percevoir l'intérêt des notions étudiées
- S'approprier le contenu de la formation

Il sera donc pertinent que les apprenants puissent réfléchir en amont de la formation à des problématiques, jeux de données ou documents susceptibles d'être utilisés en support lors de la formation.

Outil logiciel

Cette formation n'est pas strictement dédiée à un logiciel. Les exercices et les illustrations se feront généralement sous R ou sous un autre logiciel partant de l'hypothèse que celui-ci intègre les outils techniques abordés.

Dans le cas où la formation serait effectuée avec le logiciel R, une connaissance de base de ce logiciel est préconisée.

Certification

Consulter le programme sur notre [site internet](#) pour identifier si cette formation est certifiante.

Modalités pédagogiques

- Explications théoriques suivies de pratiques guidées puis de mises en autonomie.
- Exercices autonomes et réguliers pour assurer l'assimilation

Moyens et supports pédagogiques

Votre formation se déroule à distance avec :

- 1 ordinateur
- 1 connexion Internet
- 1 adresse e-mail valide
- 1 équipement audio (micro et enceintes ou casque)
- 1 Webcam (facultatif – dans l'idéal)
- 1 deuxième écran (facultatif – dans l'idéal)

Evaluation et suivi

Les objectifs pédagogiques sont évalués et suivis grâce à différentes **méthodes adaptées aux acquisitions de compétences visées**. Nos formateurs réalisent ces évaluations **tout au long de la formation**, que ce soit pendant les séances synchrones ou asynchrones. Voici une liste non exhaustive des méthodes d'évaluation pouvant être utilisées en formation :

- **Questionnaire** de vérification de connaissances (Quiz)
- Réalisation de **Travaux Pratiques** : production ou amélioration d'un fichier
- **Mises en situation** et grilles d'analyse

- **Serious Game** : jeu de rôles et analyse
- **Activités interactives** à travers l'utilisation d'une plate-forme connectée
- ...

Toutes nos formations intègrent **une auto-évaluation** via notre Extranet Stagiaire au début et à la fin de chaque action de formation. Cet outil offre à chacun la possibilité de mesurer sa progression par rapport aux objectifs pédagogiques visés et leurs atteintes.

Profil formateur

Nos formateurs sont certifiés à l'issue d'un parcours organisé par nos soins. Ils bénéficient d'un suivi de maintien et d'évolution de leurs compétences aussi bien au niveau technique que pédagogique. Chacun de nos formateurs a bénéficié d'une formation spécifique à l'animation de classe virtuelle et à l'utilisation des solutions de formation à distance du Groupe ARKESYS.

Support stagiaire

À l'issue de la formation, les exercices et travaux pratiques réalisés, leurs corrigés ainsi qu'un support de cours dématérialisé sera fourni à chaque stagiaire via son extranet stagiaire.

Accessible à tous

Cette formation est accessible à toute personne en situation de handicap. Notre référent handicap prendra contact avec les stagiaires concernés pour adapter l'animation à leurs besoins et rendre l'apprentissage accessible à tous. Enfin, nos centres de formation sont accessibles aux personnes à mobilité réduite. [En savoir plus](#)