SURAPPRENTISSAGE : MESURER L'ERREUR D'UN MODÈLE PRÉDICTIF

Durée & Format

Durée : 10 heures et 30 minutes

Format disponible : Présentiel ou Distanciel

Tarifs

Tarif public intra : Consulter notre site Internet

Tarif public inter : Consulter notre site Internet

Public visé

Personne souhaitant comprendre les phénomènes de sous et de surajustement rencontrés en Machine Learning mais aussi en statistique "classique".

Prérequis

- Cette formation n'est pas orientée sur l'apprentissage des méthodes de modélisation mais bien sur les problématiques liées à la mise en œuvre de celles-ci.
- Une connaissance des outils statistiques de base est souhaitée ainsi que la connaissance des outils de modélisation mis en œuvre

Objectifs pédagogiques

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

- Quantifier l'erreur d'un modèle prédictif qualitatif
- Quantifier l'erreur d'un modèle prédictif quantitatif
- Comprendre la notion de biais d'un modèle
- Comprendre la notion de variance d'un modèle
- Comprendre les conséquences du sous ajustement
- Comprendre les conséquences du sur ajustement
- Quantifier l'erreur d'un modèle (Quantitatif ou qualitatif)
- Mettre en œuvre des processus de validation croisée type Leave One Out et K Fold

Programme et déroulement

Jour 1 – Matin

Rappeler les méthodes de modélisation

- Que veut-on prédire
  - Variable Qualitative
  - Variable Quantitative
- Quantification de l'erreur d'un modèle quantitatif
  - Rmse,
  - Mae
  - R^2, R^2 ajusté, R^2 Prev,
  - …
- Quantification de l'erreur d'un modèle qualitatif
  - Matrice de confusion

- Accuracy
- Kappa
- Spécificité ; sensibilité
- …

- Différentes façons de prédire
- Modélisation classique
  - Régression
  - Régression logistique
- Modélisation Machine Learning
  - Knn
  - Svm,
  - Arbres de décisions
  - …

Jour 1 – Après-midi

Optimiser 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
- Matrice de confusion
- Méthode du Data Train / Data test
- Validation croisée
- Classique LOO (leave One Out)
- Validation croisée par k fold

Jour 2 – Matin

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 "reconnaitre" 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