

{tabnet} : Un package de deep-learning pour données tabulaires entièrement intégré à tidymodels

Christophe Regouby[Airbus, christophe.regouby@airbus.com]

Résumé (max 300 mots)

{tabnet} est un package d'apprentissage profond sur les données tabulaires d'après Tabnet Arik and Pfister (2019). Il est entièrement basé sur {torch} et donc sans installation de python et s'intègre dans la suite {tidymodels}. Je démontrerai l'usage du package sur un jeu de donnée d'exemple, étape par étape, depuis la gestion des valeurs manquantes, le pré-entraînement non supervisé puis l'entraînement fin sur la tâche cible, et l'interprétation du modèle. Je vous détaillerai les évolutions en cours du package et vous ferai un retour d'expérience sur le développement à cheval entre deux univers de packages {mlverse} pour le deep-learning et {tidymodels} pour la création de modèles sans faille, et je vous donnerai l'envie de contribuer à la conversion de modèles de deep-learning en R avec {torch}.

Mots-clefs : Deep-learning – Données tabulaires – torch – IA – Package

Développement

Références

Arik, Sercan Ömer, and Tomas Pfister. 2019. “TabNet: Attentive Interpretable Tabular Learning.” *CoRR* abs/1908.07442. <http://arxiv.org/abs/1908.07442>.