

Modèle linéaire généralisé et Choix de modèles

Devoir maison obligatoire

Robin Ryder et Vincent Rivoirard

Rédiger un rapport au format RMarkdown. Envoyer avant le 31 août 2019 un courriel à Robin Ryder (ryder@ceremade.dauphine.fr) avec un lien vers un dépôt github où seront présents votre fichier `.Rmd` et un fichier `.csv` contenant vos prédictions.

1 Pluie à Bâle

Le fichier `meteo.train.csv` contient des données sur les conditions météorologiques à Bâle (Suisse). Chaque ligne correspond à un jour entre 2010 et 2018. Les colonnes correspondent aux valeurs moyenne, minimale et maximale sur la journée de :

- Température ($^{\circ}C$)
- Humidité relative (pourcentage)
- Pression (hPa)
- Nébulosité (pourcentage)
- Nébulosité forte, moyenne et faible
- Vitesse (en km/h) et direction (en degrés) du vent à 10 m d'altitude, 80 m d'altitude, et à l'altitude où la pression vaut 900 hPa
- Rafales de vent à 10 m

ainsi qu'aux valeurs totales sur la journée de :

- Précipitations (mm)
- Neige (cm)
- Minutes d'ensoleillement
- Rayonnement solaire (W/m^2)

On cherche à prédire s'il pleuvra le lendemain (colonne `pluie.demain`).

Proposer et valider un modèle pour expliquer la présence de pluie le lendemain. Puis proposer une prédiction binaire pour les lendemains des journées incluses dans le fichier `meteo.test.csv`.

Source des données : MeteoBlue.