

## SY19 – A17

### TP 4 (noté) : Régression et classification – Sélection de modèles

Le but de ce TP est de traiter un problème de classification (données `tp3_clas_app.txt`) et un problème de régression (données `tp3_reg_app.txt`). Dans chaque cas, il faudra proposer le meilleur prédicteur possible, et une évaluation de la qualité de ce prédicteur (probabilité d'erreur pour le classifieur, espérance de l'erreur quadratique pour le modèle de régression).

Vous êtes libres de la méthodologie à employer. La note du TP dépendra pour 50% des performances obtenues sur chacun des deux problèmes, et pour 50% de la qualité du rendu écrit (clarté des explications; correction du français ou de l'anglais; qualité des tableaux et des figures; soin dans la présentation du rapport).

Documents à rendre au plus tard le **14 novembre à minuit** sur Moodle, sous forme d'une archive zip (impérativement) contenant exactement **quatre** fichiers, où X est le nom de votre groupe (lettre entre A et V) :

- Rapport écrit au format pdf réalisé avec un *notebook* RStudio, en français ou en anglais, maximum 8 pages (nom de fichier : `rapport_X.pdf`)
- La fonction codant le classifieur (nom de fichier : `classifieur_X.R`, nom de la fonction : `classifieur`).
- La fonction codant le régresseur (nom de fichier : `regresseur_X.R`, nom de la fonction `regresseur`).
- Un fichier de données R contenant l'environnement nécessaire à l'exécution des deux fonctions ci-dessus (nom de fichier : `env_X.Rdata`).

Les fonctions `classifieur` et `regresseur` devront avoir comme seul argument le *data frame* contenant les données de test, et renvoyer en sortie un vecteur contenant les prédictions pour les données de test. Toute erreur dans l'exécution du code entraînera l'absence de point pour la partie correspondante.

Un exemple de fichier `classifieur_X.R` :

```
classifieur <- function(dataset) {  
  # Chargement des données construites lors de l'apprentissage (si besoin)  
  load("env_X.Rdata")  
}
```

```

# Mon algorithme qui renvoie les prédictions sur le jeu de données
# `dataset` fourni en argument.
# ...

return(predictions)
}

```

Un exemple de fichier `regresseur_X.R` :

```

regresseur <- function(dataset) {
  # Chargement des données construites lors de l'apprentissage (si besoin)
  load("env_X.Rdata")

  # Mon algorithme qui renvoie les prédictions sur le jeu de données
  # `dataset` fourni en argument.
  # ...

  return(predictions)
}

```