

## Suivi du projet + Ordre du jour

**Prochain rendez-vous :** Vendredi ?, Université Paul Valéry, bâtiment B.

### **Travail effectué :**

- Matrice X complétée, log10 appliqué.
- Préparation du prochain rapport sur les données et leur exploration à rendre pour 20h.
- En cours : Créer les classes : up, down et neutre en considérant un seuil à 2 pour le up (resp. 1/2 pour les down) pour le ratio des comptages de read ARNm ("fold change").
- En cours : Analyse exploratoire des données.

### **Travail à effectuer prochainement :**

- Classe Y : Répartir en classe les résultats. up, down et neutre en considérant un seuil à 2 pour le up (resp. 1/2 pour les down) pour le ratio des comptages de read ARNm ("fold change"). + Regarder la distribution.
- Regarder les liens entre les variables, à l'intérieur de la matrice X avec le cours de régression linéaire logistique (régression deux à deux, conditionnement de X transpose X).
- Classification hiérarchique. Matrice de distance deux à deux en entrée (cela peut être la corrélation).  
ACP + Clustering avec X. Puis chaque colonne de Y. La classification commence quand on fait cela.  
1er modèle) Régression logistique. Coder un forward. Package R qui fait ça : LASSO (Norme L1), Ridge (Norme L2).
- Découvrir GLMnet, tester sur les données.
- PRÉVOIR DES CRITÈRES D'ÉVALUATION DU MODÈLE.
- Calculer la précision et le rappel.
- Représentation sous forme de heatmap, cases coloriées.

### **Ce que nous aimerions aborder lors de la réunion :** Liste non ordonnée.

1. Définition date/heure du prochain rendez-vous.
2. Revue de la matrice X + discussions sur la matrice Y + données manquantes (Na)
3. Rapport sur les données et leur exploration.
4. Package GLMnet.
5. Autres questions suite au travail perso.