

Planification et gestion de projet

Organisation du groupe :

Le groupe est composé de 3 membres : Matis Breillad, Florian Dubois, Alvin Vedel. Alvin occupe le rôle de chef de projet ce qui correspond à la planification des tâches, la communication avec l'encadrant et le bon déroulement du projet.

Nos points forts se complétant plutôt bien, Matis utilisera R pour l'analyse de donnée puis passera sur python pour l'implémentation de modèles, Florian et Alvin à l'aise en python l'utiliseront pour la manipulation de donnée et établir des premiers modèles.

Un github a rapidement été mis en place pour l'échange de fichiers, les communications relatives au projet s'effectue sur discord permettant un archivage et une consultation rapide.

Les rendez-vous avec le commanditaire sont planifiés par mail universitaire avec chacun des membres en copie pour une bonne circulation de l'information. Nous avons jugé avec M. Servajean la nécessité d'un rendez-vous par période de cours et parfois plus si besoin et en cas de questions.

Planification des tâches :

Les tâches principales du premier semestre sont :

- L'analyse des données
- L'implémentation de modèles prédictifs

Les sous-tâches qui en découlent sont les suivantes :

Analyse de données :

Création de visualisation sur les variables quantitatives, visualisations cartographiques sur la répartition des espèces. Analyse en réduction de dimension sur les données abiotiques (ACP) et représentation de rasters environnementaux.

Implémentation de modèles :

Création de modèles constants (aléatoires et plus présents), K voisins spatiaux et abiotiques. Implémentation de modèles utilisant les données de présences uniquement. Forêt aléatoire sur les données abiotiques et SDM.

Réparties de la façon suivante :

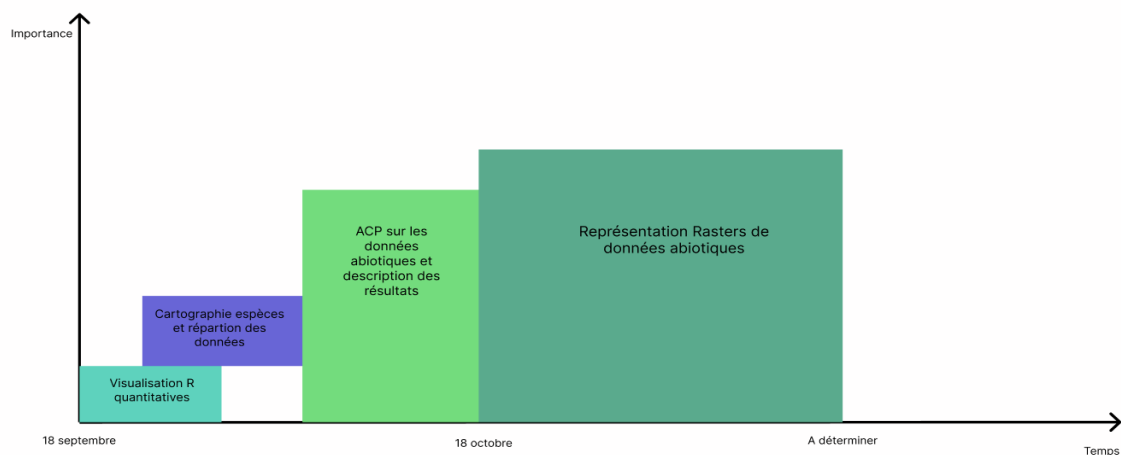


Figure 1 : Gantt tâches d'analyse

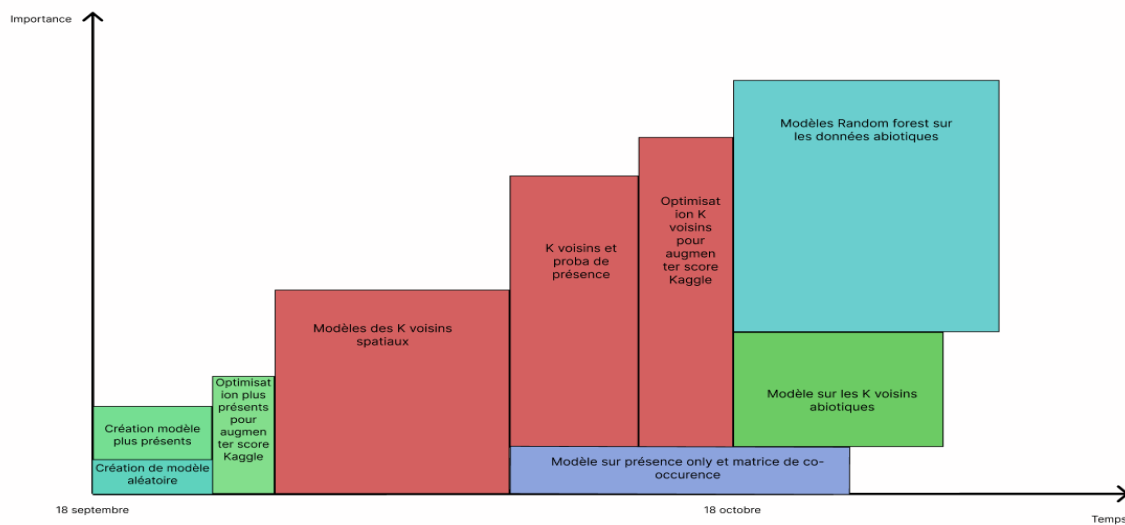


Figure 2: Gantt tâches modèles

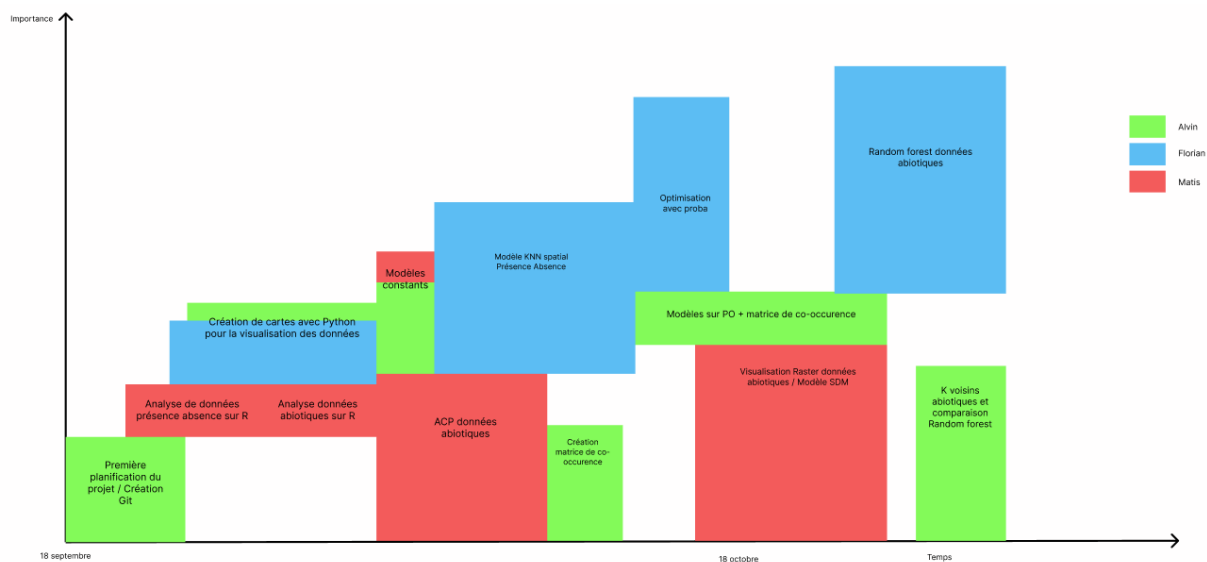


Figure 3: Gantt de la répartition globale

Choix de l'outil de gestion de projet :

Les outils de gestion de projet testés sont les suivants : Excel, Trello et Github. Un premier planning a été réalisé sur Excel permettant de se représenter le déroulement du premier semestre, il ne permet cependant pas une bonne gestion et répartition des tâches. Un projet Github et Trello ont plus ou moins les mêmes fonctionnalités, permettant d'attribuer des tâches et d'en suivre l'avancement : catégorisé par « à faire », « en cours » et « terminé ». Trello possédait une interface utilisateur plus agréable et d'autres fonctionnalités comme la création de listes supplémentaires cependant nous avons opté pour un projet Github. L'outil permet un bon suivi des tâches mais enregistre également les issues liées aux différents dépôts des membres du groupe. Cela nous permet d'indiquer la qualité de nos fichiers et les modifications nécessaires tout en centralisant l'information, le Github contenant donc l'ensemble de notre TER et le suivi de notre projet.

Projet TER						
View 1 + New View						
Filter by keyword or by field						
Title	Assignees	Status	Description			
7 Rédaction du rendu 1	AlvinVedel, Florianj...	Done	Rédaction collaborative du premier rapport			
8 ACP exploratoire	Matis24	Done	ACP sous R pour établir lien entre variables			
9 Modèles K-voisins	Florianjava	Done	Réalisation en python de modèles plus complexes pour amélioration score Kaggle			
10 Outils de gestion de projet	AlvinVedel	Done	Mettre en place les outils de gestion de projet : Trello etc.			
11 Réunion Octobre avec commanditaire	AlvinVedel, Florianj...	In Progress	Planifier réunion et s'y rendre et compte rendu			
12 Problème K-voisins #2	AlvinVedel, Florianj...	Done				
13 Amélioration ACP #3	AlvinVedel, Florianj...	Done				
14 Amélioration K-voisins	Florianjava	Done	Amélioration en sélectionnant ceux présents au moins dans X voisins			
15 Création matrice co-occurrence	AlvinVedel	Done	matrice de co occurrence - densité de probabilité			
16 Analyse et création matrice des k-voisins / proba	Florianjava and Mat...	In Progress	Matrice avec différentes valeurs de k et différentes proba pour k voisins			
17 Amélioration meilleur K-voisin avec matrice de co-occurrence						
18 Modèles Noyaux	Florianjava	Todo	Réalisation en python de modèles plus complexes pour amélioration score Kaggle			
19 Rédaction rendu 2	AlvinVedel, Florianj...	Todo	Rédaction collaborative du second rapport			
20 Intégration des Données Présence / Only	AlvinVedel	In Progress	Modèle Presence Only et matrice de co-occurrence			
21 Développement Interface Web - Back-End	AlvinVedel	Todo	Développement backend du site web			
22 K-voisins "abiotiques"		Todo	Rechercher les zones pas les plus proches géographiquement mais selon composition...			
23 Modele Clustering Données Abiotiques	Matis24	Todo	Projection des données abiotique de test dans les clusters obtenus via ACP			

Figure 4: Illustration de la gestion de projet Git

Stratégie et jalons :

La stratégie consiste d'abord en l'analyse de données puis l'implémentation de modèles. Les modèles se complexifieront au fil du semestre en englobant davantage de données. Le principal indicateur de notre avancement est le score Kaggle qui juge de la précision du modèle. Ce score compris entre 0 et 1 va augmenter par palier, le meilleur score actuellement enregistré sur le Kaggle est 0.276 en public et 0.2705 en privé. Le score public correspond à la précision du modèle sur le jeu de donnée de validation, c'est celui-ci que nous essayons de maximiser. Le score privé correspond quant à lui à la précision du modèle sur le jeu de donnée de test, il est censé être inconnu donc il ne faut pas tester nos paramètres pour l'améliorer.