Plusieurs aspects négatifs nous ont empêchés de mieux effectuer ce travail.

Avant le commencement de la programmation :

Trop de temps a été passé à tenter d'installer correctement les librairies de ROCR sur OSX n’utilisant pas R studio le temps à mettre en place l'environnement de travail avec les librairies nécessaire a été problématique. De plus, beaucoup de temps a été mi à la compréhension de la tache qui était à effectuer.

Lors de la programmation :

le plus gros défi qu'on a eu c’était que notre compression du problème était très fragmentée.

On comprenait plusieurs morceaux du puzzle, mais on ne savait pas dans quel ordre les placers ni si on les avait tous.

Donc explication de ce que nous avons fait, en premier lieu nous avons chargé en mémoire le fichier CSV.

Puis nous avons séparé le fichier en deux ensembles un de test et l’autre d’entrainement;

Ensuite on construit notre arbre de récursion avec RPART avec les données d’entrainement. Nous trouvons alors une valeur initiale pour l'élagage.

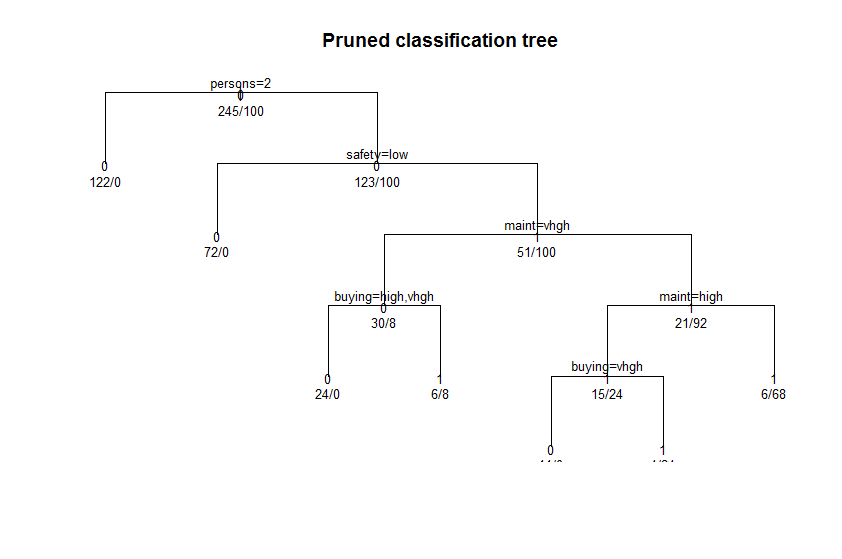
Pour valider notre arbre c’est là qu'on a eu plus de problèmes on construit des courbes ROC. Pour produire ces courbes ROC on utilise la fonction Predict avec en paramètre notre arbre de et l’ensemble des données de test. Nous varions ensuite le critère d'erreur pour avoir la meilleure courbe ROC possible.

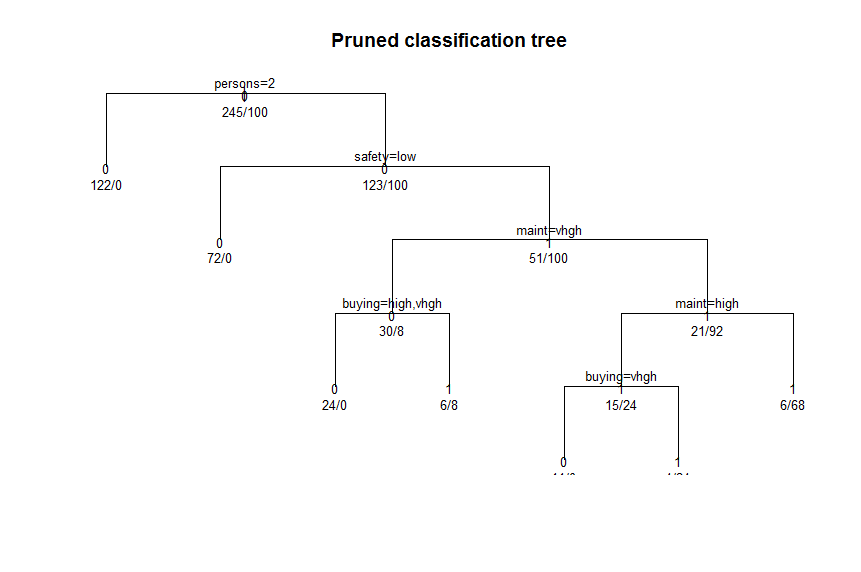
Après avoir programmé :

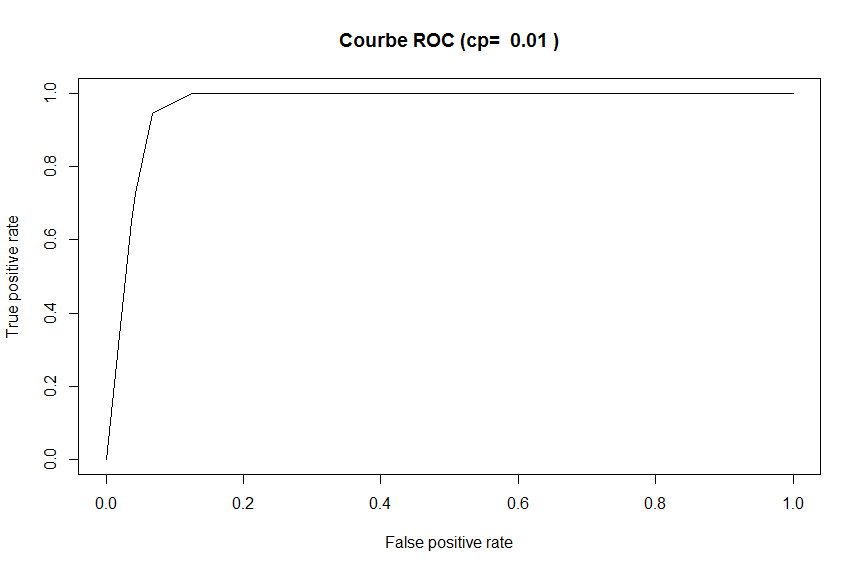
Comme expliqué plus haut, notre compression est très fragmentée, on a fini par bien comprendre comment fonctionne un arbre de récursion, mais pour ce qui est de l’élagage je ne peux pas dire que je comprends le pourquoi mathématique et algorithmique c’est plus jouer avec les chiffres avec la fonction. De plus notre élagage ne semblait avoir aucun impact sur notre arbre.

Pour ce qui est des courbes ROC, nous croyions avoir compris la base, mais ça va demander plus de temps d’étude pour bien comprendre les subtilités.

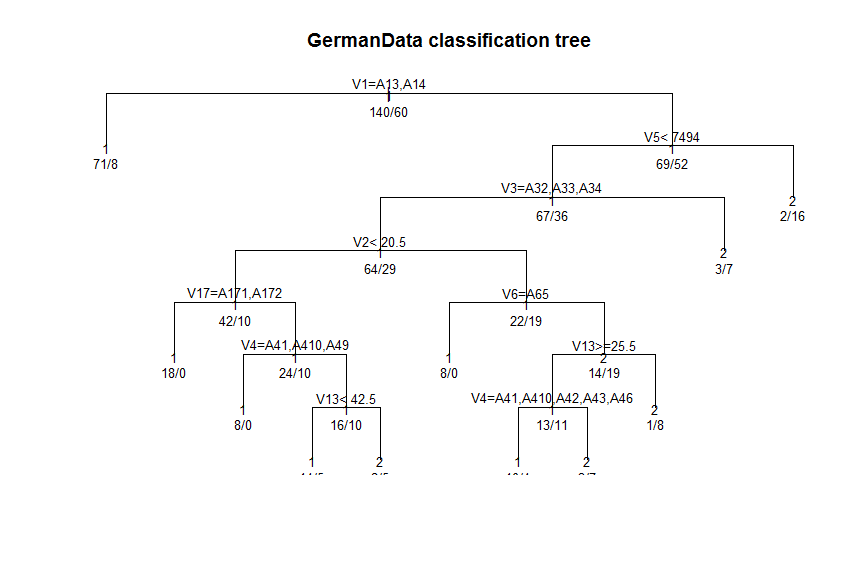
**Graphiques générés pour CarData:**

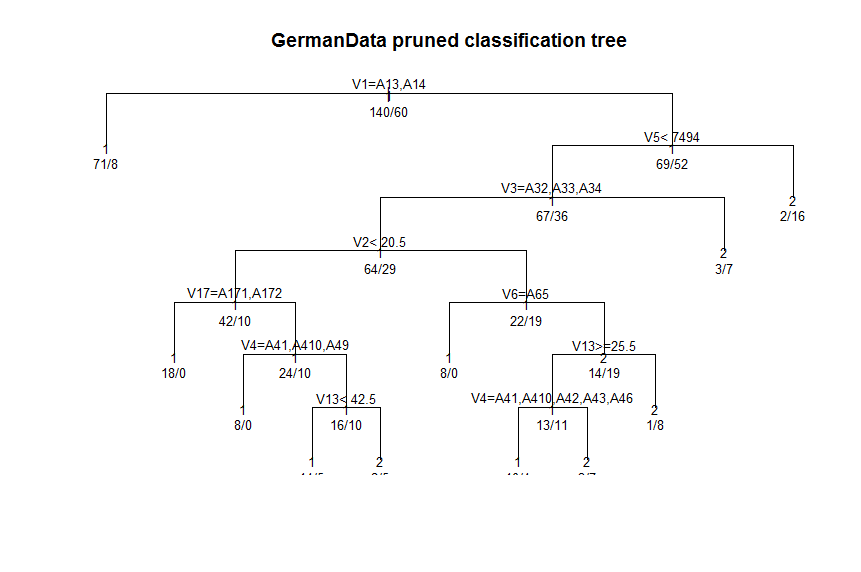
****

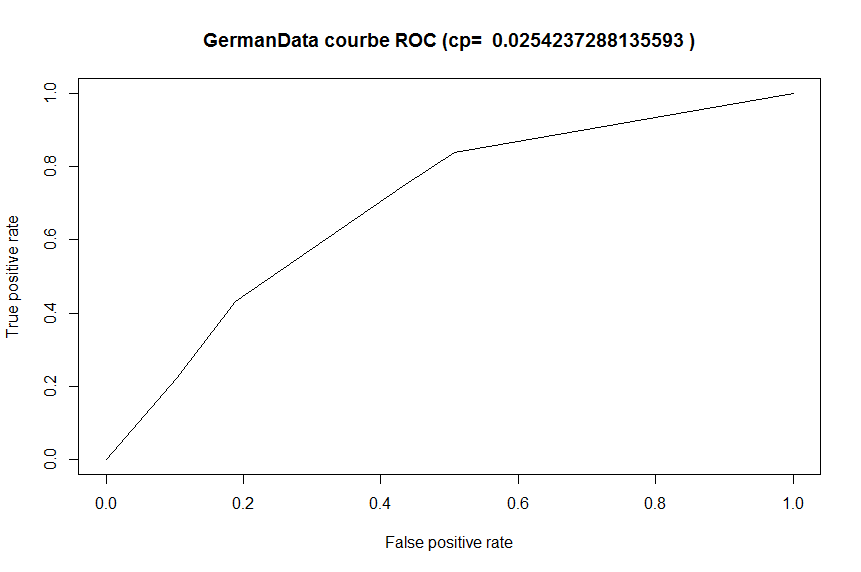
****

****

**Graphiques générés pour GermanData:**

****

****

****