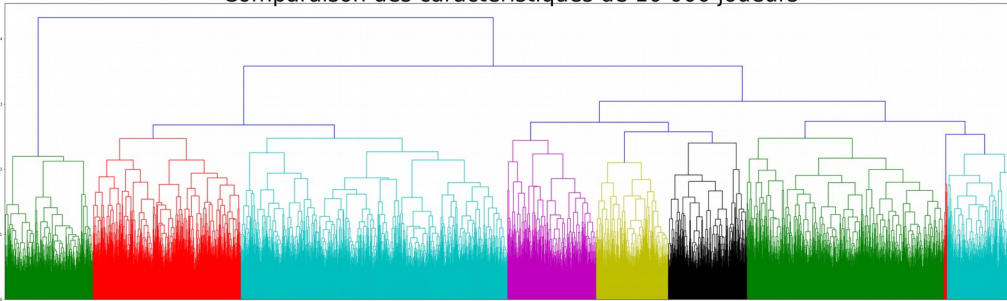


Datascience pour le football

Apprentissage non supervisé:

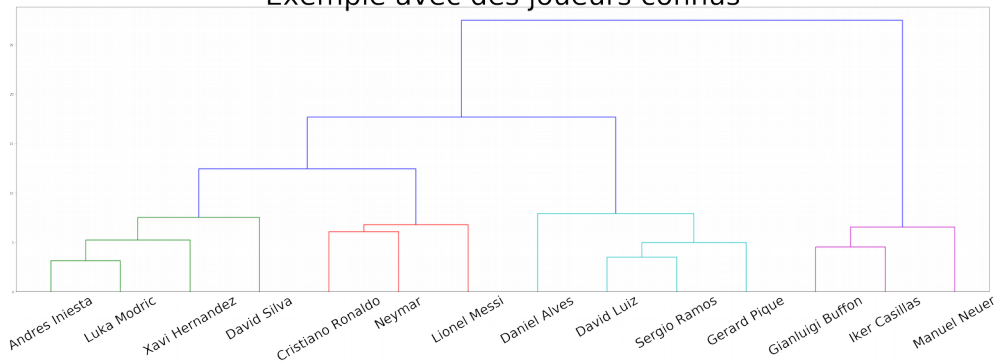
Peut on retrouver les postes des joueurs, à partir de leurs caractéristiques ?

Comparaison des caractéristiques de 10 000 joueurs



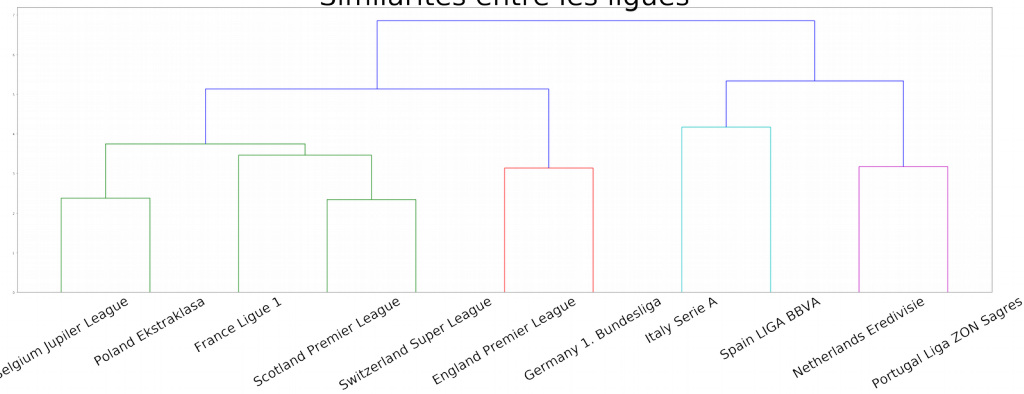
Ce dendrogramme montre qu'il existe bien des groupes de joueurs similaires, mais n'est pas très lisible dû au grand nombre d'individus, voici un exemple plus clair.

Exemple avec des joueurs connus



De la même façon, peut on comparer les leagues selon leur niveau moyen ?

Similarités entre les leagues



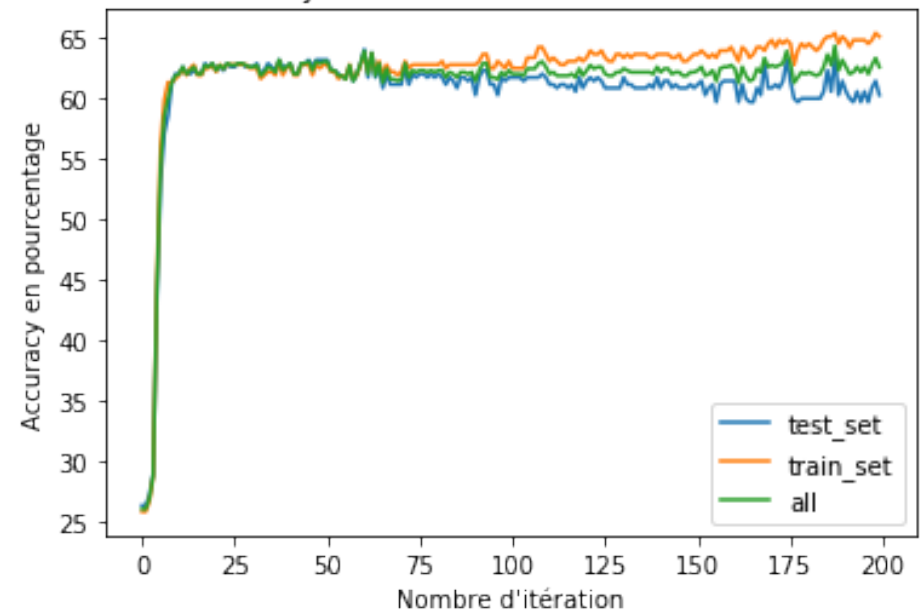
Apprentissage supervisé:

Peut on prédire le résultat d'un match en fonction des caractéristiques des deux équipes ?

Pour répondre à cette question nous avons utilisé comme données la liste des matchs et nous avons formé notre base d'apprentissage avec pour attributs les caractéristiques des deux équipes et leur ratio de victoires, et comme label la victoire de l'équipe jouant à domicile, ou non.

Modèle prédictif avec un perceptron avec kernel polynomial.

Accuracy en fonction du nombre d'itérations



Modèle prédictif avec un arbre de décision

Avec un arbre de décision, nous trouvons une accuracy de 70% avec le test set.