

# Comparaison qualitative du coup droit au tennis entre le jugement d'un expert et des algorithmes d'apprentissage.

Fanny Mauron

Master thesis in Sport Science

De nombreuses nouvelles avancées dans le domaine informatique ont permis de développer l'utilisation d'algorithmes dans des domaines de la vie quotidienne. Le but de ce travail est de mettre en relation le Machine Learning et le domaine sportif. Il s'agit de définir s'il est possible de prédire un jugement qualitatif humain par le biais d'un apprentissage machine supervisé. Comme mouvement de base, nous avons choisi le coup droit au tennis.

559 coups droits exécutés par des sujets hommes experts et débutants au tennis ont été capturés en 3D. Ces mouvements ont ensuite été jugés qualitativement par les experts qui leur ont attribué une note de 0 à 9. Nous avons ainsi obtenu une base de données d'apprentissage pour le Machine Learning. Deux algorithmes de classification et de régression ont été choisis, les k-plus proches voisins et les arbres de décision. Nous avons également testé une méthode de régression linéaire et polynomiale ainsi qu'une méthode de deep-learning et de boosting du gradient. Tous ces algorithmes ont prédit les notes pour les données test et une comparaison a été faite avec les vraies valeurs. Nous avons calculé le score de prédiction pour les méthodes de classification et le coefficient de détermination  $R^2$  pour les méthodes de régression. L'erreur quadratique moyenne ainsi que l'erreur absolue moyenne ont également été calculées afin de pouvoir comparer les différentes méthodes d'apprentissage.

Les méthodes de régression sont meilleures que les méthodes de classification. Par-dessus tout, la méthode de boosting du gradient est la méthode qui prédit les notes avec la plus grande justesse. C'est la seule méthode qui obtient une erreur moyenne absolue plus petite que 0,7 et un coefficient  $R^2$  de 0.78. Nous avons pu définir qu'il existe suffisamment de cohérence dans les différents jugements des experts pour que l'algorithme d'apprentissage puisse généraliser les associations et prédire avec une précision acceptable les notes de nouveaux mouvements. Cependant, les résultats étant très diversifiés, nous n'avons pas pu définir précisément les caractéristiques importantes qui ont permises aux experts de distinguer qualitativement les mouvements entre eux. Plusieurs problèmes ont été relevés dans le jeu de données d'apprentissage tels que le manque d'homogénéité dans la répartition des données dans les classes et la rigueur dans le jugement des experts.

Le Machine Learning offre de très larges possibilités d'utilisations dans le domaine sportif. La récolte des données et la préparation des données d'apprentissage sont des étapes très importantes à ne pas sous-estimer pour que les résultats soient précis et explicables.

**Superviseur :** Prof. Jean-Pierre Bresciani