



Université Claude Bernard



Lyon 1

DIPLÔME NATIONAL DE DOCTORAT

(Arrêté du 25 mai 2016)

Date de la soutenance : **15 novembre 2019**

Nom de famille et prénom de l'auteur : **SCHMUTZ Amandine**

Titre de la thèse : « *Contributions à l'analyse de données fonctionnelles multivariées, application à l'étude de la locomotion du cheval de sport* ».



Résumé

Avec l'essor des objets connectés pour fournir un suivi systématique, objectif et fiable aux sportifs et à leur entraîneur, de plus en plus de paramètres sont collectés pour un même individu. Une alternative aux méthodes d'évaluation en laboratoire est l'utilisation de capteurs inertiels qui permettent de suivre la performance sans l'entraver, sans limite d'espace et sans procédure d'initialisation fastidieuse. Les données collectées par ces capteurs peuvent être vues comme des données fonctionnelles multivariées : se sont des entités quantitatives évoluant au cours du temps de façon simultanée pour un même individu statistique.

Cette thèse a pour objectif de chercher des paramètres d'analyse de la locomotion du cheval athlète à l'aide d'un capteur positionné dans la selle. Cet objet connecté (centrale inertielle, IMU) pour le secteur équestre permet de collecter l'accélération et la vitesse angulaire au cours du temps, dans les trois directions de l'espace et selon une fréquence d'échantillonnage de 100 Hz. Une base de données a ainsi été constituée rassemblant 3221 foulées de galop, collectées en ligne droite et en courbe et issues de 58 chevaux de sauts d'obstacles de niveaux et d'âges variés.

Nous avons restreint notre travail à la prédiction de trois paramètres : la vitesse par foulée, la longueur de foulée et la qualité de saut. Pour répondre aux deux premiers objectifs nous avons développé une méthode de clustering fonctionnelle multivariée permettant de diviser notre base de données en sous-groupes plus homogènes du point de vue des signaux collectés. Cette méthode permet de caractériser chaque groupe par son profil moyen, facilitant leur compréhension et leur interprétation. Mais, contre toute attente, ce modèle de clustering n'a pas permis d'améliorer les résultats de prédiction de vitesse, les Support Vector Machine (SVM) restant le modèle ayant le pourcentage d'erreur inférieur à 0.6 m/s le plus faible. Il en est de même pour la longueur de foulée où une précision de 20 cm est atteinte grâce aux SVM. Ces résultats peuvent s'expliquer par le fait que notre base de données est composée uniquement de 58 chevaux, ce qui est un nombre d'individus très faible pour du clustering.

Nous avons restreint notre travail à la prédiction de trois paramètres : la vitesse par foulée, la longueur de foulée et la qualité de saut. Pour répondre aux deux premiers objectifs nous avons développé une méthode de clustering fonctionnelle multivariée permettant de diviser notre base de données en sous-groupes plus homogènes du point de vue des signaux collectés. Cette méthode permet de caractériser chaque groupe par son profil moyen, facilitant leur compréhension et leur interprétation. Mais, contre toute attente, ce modèle de clustering n'a pas permis d'améliorer les résultats de prédiction de vitesse, les Support Vector Machine (SVM) restant le modèle ayant le pourcentage

d'erreur inférieur à 0.6 m/s le plus faible. Il en est de même pour la longueur de foulée où une précision de 20 cm est atteinte grâce aux SVM. Ces résultats peuvent s'expliquer par le fait que notre base de données est composée uniquement de 58 chevaux, ce qui est un nombre d'individus très faible pour du clustering.

Nous avons ensuite étendu cette méthode au co-clustering de courbes fonctionnelles multivariées afin de faciliter la fouille des données collectées pour un même cheval au cours du temps. Cette méthode pourrait permettre de détecter et prévenir d'éventuels troubles locomoteurs, principale source d'arrêt du cheval de saut d'obstacle.

Pour finir, nous avons investigué les liens entre qualité du saut et signaux collectés par l'IMU. Nos premiers résultats montrent que les signaux collectés par la selle seuls ne suffisent pas à différencier finement la qualité du saut d'obstacle. Un apport d'information supplémentaire sera nécessaire, à l'aide d'autres capteurs complémentaires par exemple ou encore en étoffant la base de données de façon à avoir un panel de chevaux et de profils de sauts plus variés.