



03 JUIN

🕒 09h00 à 11h00

📍 ENS Paris-Saclay

THÈSES ET HDR

Mounir ATIQ : Monitoring des personnes âgées à partir de modèles d'arbres de décision...

Direction : Nicolas Vayatis et Mathilde Mougeot

Soutenance : le 03/06/22 en 1B36

📅 AJOUTER AU
CALENDRIER

Titre

Monitoring des personnes âgées à partir de modèles d'arbres de décision en condition de variation de domaines et sous contraintes de capacités de calcul

Résumé

Cette thèse porte sur le monitoring de personnes âgées en maisons de retraite et EHPAD, à partir d'un capteur sol piézo-électrique connecté à un système embarqué. Après avoir présenté les particularités du signal piézo-électrique ainsi que sa conversion en séries temporelles, nous expliquons comment les différentes technologies de mesures affectent le signal original et présentons les différents environnements de collecte de données. Plusieurs études expérimentales sur les forêts aléatoires sont réalisées et montrent la faisabilité de plusieurs tâches de détection, ainsi que la pertinence des différentes représentation du signal et features associées.

Cependant pour être déployés de manière industrielle ces modèles doivent respecter deux contraintes majeures. Premièrement ils doivent être confrontés aux données réelles et s'adapter à la variabilité des installations et des patients. Dans cette optique, des méthodes d'apprentissage par transfert sont étudiées pour l'intégration de nouvelles données obtenues en conditions réelles dans un modèle préalablement entraîné sur des données simulées en environnement contrôlé. Nous étudions les effets néfastes liés au déséquilibre de classes sur l'apprentissage par transfert et proposons, pour palier à ce problème, des adaptations de méthodes existantes sur les arbres de décision. A partir de ces adaptations nous développons un algorithme robuste de transfert sur les forêts aléatoires utile à la fois pour gérer le déséquilibre de classes et pour interpréter les relations entre les différents domaines de données.

En outre ces modèles doivent tourner en temps réel dans un système embarqué aux capacités de calcul réduites. Prendre en compte ce type de contraintes lors de l'apprentissage statistique correspond au domaine de recherche appelé « apprentissage budgété », qui est particulièrement actif ces dernières années. Nous définissons formellement le temps de prédiction pour les arbres de décision et forêts aléatoires prenant en compte à la fois le coût d'acquisition des features et le coût d'évaluation du modèle. Nous proposons un algorithme génétique pour résoudre le problème du budget sur le temps de prédiction qui permet de passer d'un modèle déjà entraîné sans limite de complexité à un modèle simplifié qui respecte cette contrainte de temps de calcul. Cet algorithme tire parti d'un pruning aléatoire et de la notion d'équivalence entre les arbres de décision, à savoir lorsque deux modèles représentent la même fonction de décision mais diffèrent par leur structure. Cela permet ainsi de favoriser la réduction du temps de prédiction des modèles en exploitant la variété structurelle des arbres de décision.