**Rapport Final**

**Apprentissage Profond**

**8INF892**

**Kévin Bouchard**

**Université du Québec à Chicoutimi**

**Département d’Informatique**

**Jean-Philippe Larouche**

**LARJ07089501**

**29/04/2024**

# Introduction

# Contexte

Ce projet s’inscrit dans le domaine de la reconnaissance d’activité humaine (RAH), la tendance des LLMs et comme première étape d’un plus grand projet de recherche. L’approche actuelle en RAH est d’utiliser des modèles d’apprentissage profond qui son entrainer spécifiquement pour reconnaitre une ou un petit ensemble d’activités spécifiques. De plus, l’entrainement de ces modèles nécessitent beaucoup de données de qualité. Cela demande un grand effort d’annotation, nettoyage, de structuration et de collecte de la part de chercheurs.

Dans le contexte du cours d’apprentissage profond (8INF892) à l’UQAC, nous devions entreprendre un projet avec l’objectif d’avancer nos connaissances du domaine. Dans cette optique, nous avons pris la décision d’entreprendre un projet sur les *Large Lanugage Models* (LLM) dans le contexte de la reconnaissance d’activités humaines. Plus spécifiquement, nous avons fait une étude comparative sur différents LLMs populaire en ce qui concerne leurs capacités à interpréter des données brutes provenant de différents capteurs.

# Objectifs

Tel que mentionné, ce projet est une première étape d’un plus grand projet de recherche qui à l’objectif de déterminer si les LLMs peuvent reconnaitre les activités performées par un humain en utilisant les données brutes provenant des capteurs situés au sein d’un habitat intelligent. Dans ce projet, nous commençons par évaluer les capacités de différents LLMs à traiter des données de capteurs simples dans le contexte de la RAH. Nous nous sommes inspirés de l’article *HARGPT : Are LLMs Zero-Shot Human Activity Recognizers?* de Ji et al. [1] qui effectue une analyse similaire avec GPT4. L’objectif de ce projet est de déterminer les capacités des différents LLMs populaire sur le marché a fin d’être en mesure d’établir ceux qui seront utilisés pour les prochaines étapes du plus grand projet. Comme objectifs secondaires, nous voulons également :

* Déterminer la meilleure façon de structurer l’invite de texte
* Évaluer comment passer les données au modèle
* Apprendre à utiliser LM Studio pour faire fonctionner un LLM en local
* Évaluer la capacité de raisonnement des LLMs (pas *zero-shot*)

# Chercheurs

Ce projet de recherche fût effectué seul par Jean-Philippe Larouche, étudiant à la maîtrise de recherche en Informatique.

# Méthodologie

# Données

Les données représentent une des composantes les plus importantes de ce projet. Tel que mentionnée dans l’introduction, de bonnes données sont essentielles afin d’être en mesure de faire de la RAH. Initialement, nous avions prévus utilisés des données provenant d’un environnement « complexe », soit les données provenant du laboratoire de domotique du LIARA. Cependant, nous avons rapidement réalisé l’ampleur de la tâche de structuration de ces données afin de pouvoir les utilisés. Étant donné que notre objectif est d’évaluer les modèles de type LLM et non de faire de la structuration de donnée complexe, nous avons alors décidé de prendre l’ensemble de donnée Capture24 [2].

L’ensemble de donnée Capture24 contient des données d’un capteur de poignet porter par 151 participants pendant plus de 24 heures. Cela nous apporte une grande variété d’activités quotidiennes performer par les humains tel que marcher avec son chien, faire la vaisselle, dormir, écouter le TV, etc.

# Tests

# Métriques

# Invite de texte

# Résultats

# Conclusion

# Références

[1] Ji S., Zheng X., Wu C. 2024. HARGPT: Are LLMs Zero-Shot Human Activity Recognizers. arXiv:2403.02727 [cs.CL]. <https://doi.org/10.48550/arXiv.2403.02727>

[2] Chan Chang, S., Walmsley, R., Gershuny, J., Harms, T., Thomas, E., Milton, K., Kelly, P., Foster, C., Wong, A., Gray, N., Haque, S., Hollowell, S., & Doherty, A. (2021). Capture-24: Activity tracker dataset for human activity recognition. University of Oxford.