

Rapport de Sprint 2

Apprentissage Profond

8INF892

Kévin Bouchard

Université du Québec à Chicoutimi

Département d'Informatique

Jean-Philippe Larouche

LARJ07089501

12/04/2024

Travail réalisé

Dans le cadre du deuxième sprint, moins de travail fût effectué que prévu étant donné des facteurs externes (i.e. déménagement). Voici le travail qui fut réalisé par le seul membre de l'équipe, Jean-Philippe Larouche, pendant le sprint:

- Données à analyser ont été finalisées
 - Structure des données
 - Type de données
 - Quantité et séparation en groupe selon le test
- Les tests à effectuer sont établis
- Les métriques afin de pouvoir comparer les différents modèles sont établis

Alors qu'il était prévu de mettre en place le code et la structure nécessaire afin de pouvoir effectuer les tests sur les différents modèles, cela n'a pas été fait dû à un manque de temps. Le projet est définitivement en retard sur son avancement. Dans cet optique, nous avons décidé de ne pas effectuer de tests sur Claude 3 car il n'est pas aussi facilement accessible que les autres et nous gardons tout de même deux modèles grand marché utilisables à travers un API (Chat GPT 3.5 et Gemini) et deux via LMStudio (Starling LM et Llama 2).

Bilan

Données finalisées

Alors que nous avions initialement l'intention d'utiliser des données complexes de différents capteurs, nous nous sommes rendus à l'évidence que le traitement de ces données était une tâche énorme et complexe. Cela dépassait largement le temps disponible pour ce projet. Nous avons alors décidé d'utiliser des données déjà disponibles et structurées qui furent utilisées dans d'autres travaux.

- <http://archive.ics.uci.edu/dataset/240/human+activity+recognition+using+smartphones>
- <https://www.v7labs.com/open-datasets/dahlia>

Nous sommes confiants que cet ensemble de données suffiront dans notre contexte et pour notre objectif.

Tests à effectuer établis

Les tests à effectuer sur les modèles sont les suivants :

- *Zero-Shot* sans modifications
 - Un prompt initial expliquant l'objectif (HAR) puis on passe les données pour voir le résultat
- *Zero-shot* avec explication des données

- Un prompt initial expliquant l'objectif (HAR) mais on inclut des descriptifs (capteurs, types, etc.) des données. Puis, on passe les données pour voir le résultat
- *Zero-shot* avec une liste des activités possibles
 - Un prompt initial expliquant l'objectif (HAR) et on inclut les activités possibles avec les données fournies
- Explication des données
 - Ici l'objectif est de fournir des données au modèles et simplement demander son interprétation de ce que c'est.
- Retirer des séquences de données
 - Ce test se base sur ceux mentionnés précédemment mais certaines séquences de données seront retirées pour évaluer la résilience du modèle
- Données extrêmes
 - Ce test se base sur ceux mentionnés précédemment mais certaines données dans l'extrême (incohérente) seront insérées pour évaluer la résilience du modèle

Métriques de comparaisons

Nous allons avoir quatre métriques de comparaisons afin d'évaluer les modèles, qui peuvent contenir des sous métriques d'évaluation:

- Capacité à comprendre les valeurs numériques
 - Est-ce que le modèle parvient à traduire un ensemble de valeurs numériques en une activité performée par un humain
 - Quel type de donnée numérique le modèle comprend-il
- Résilience aux variations de données
 - Est-ce le modèle est toujours en mesure de reconnaître les activités lorsque on enlève de séquence de donnée ou on introduit des fluctuations de données dans les extrêmes
- Reconnaissance de l'activité
 - Est-ce que le modèle identifie adéquatement l'activité
- Aide nécessaire
 - Cette catégorie inclut toute intervention humaine nécessaire pour aider le modèle à avoir une performance satisfaisante. Par exemple, avons-nous besoin de lui fournir une liste des activités possibles pour qu'il puisse les reconnaître?

Planification Sprint 3

Pour le troisième sprint, qui termine le 29 avril 2024 et qui est le sprint final, les objectifs sont les suivants :

- Mettre en place la structure, code et autres outils nécessaires afin d'effectuer les tests sur les modèles sélectionnés.
- Évaluer et comparer les résultats entre les modèles
- Écrire le rapport final