

Stage de Master recherche (ou école d'ingénieurs)

Amélioration des modèles de langage de grande taille multimodaux (MLLM)

Période de stage : mars 2024 à juillet 2023 (5 mois)

Encadrants : Olivier CASPARY (olivier.caspary@univ-lorraine.fr) / Fahed Abdallah (fahed.abdallah@univ-lorraine.fr)

Laboratoire de recherche : LCOMS (<https://lcoms.univ-lorraine.fr/>)

Description :

Titre : Amélioration des modèles de langage de grande taille multimodaux par des techniques d'ancrage : un cadre d'évaluation et d'optimisation complet

Le sujet de master proposé vise à faire avancer le domaine des modèles de langage de grande taille multimodaux (MLLM) en se concentrant sur les techniques d'ancrage et leur impact sur la performance du modèle. Le projet s'appuiera sur le récent développement du benchmark d'évaluation MLLM (MME), qui répond au besoin d'une évaluation complète des capacités de perception et de cognition des MLLMs.

L'un des objectifs de ce projet est d'explorer et de mettre en œuvre diverses techniques d'ancrage pour améliorer les capacités perceptuelles et cognitives des MLLMs. Les techniques d'ancrage impliquent de lier les éléments linguistiques à des entités du monde réel, offrant ainsi une base plus solide pour la compréhension du modèle.

Nous devrions répondre aux préoccupations concernant la fuite de données en concevant manuellement des paires d'instructions et de réponses pour les sous-tâches MME et en développant une approche systématique pour créer des annotations concises et efficaces, garantissant une comparaison équitable entre différents MLLMs sans nécessiter une ingénierie de prompt étendue. Nous visons également à explorer les modèles et à identifier les domaines d'amélioration dans les modèles existants, fournissant des perspectives sur les directions potentielles pour l'optimisation ultérieure du modèle.

<https://arxiv.org/pdf/2306.13549.pdf>

<https://github.com/BradyFU/Awesome-Multimodal-Large-Language-Models/tree/Evaluation>

<https://cloud.google.com/blog/products/ai-machine-learning/how-to-use-grounding-for-your-llms-with-text-embeddings?hl=en>

Keywords : MLLM, techniques d'ancrage, Python.

Conditions :

Lieu de stage et employeur : LCOMS - IUT de Saint-Dié, 88100 Saint-Dié-des-Vosges, France

Rémunération : gratification légale.

Profil attendu : étudiant(e) en master/école d'ingénieurs motivé(e), volontaire, avec des compétences en programmation Python. Des connaissances en modèles de langage seraient un plus.

Research Master's Internship (or engineering school internship) "Enhancement of Multimodal Large Language Models (MLLM)"

Internship period: April 2023 to August 2023 (5 months)

Supervisors: Olivier CASPARY (olivier.caspary@univ-lorraine.fr) / Fahed Abdallah (fahed.abdallah@univ-lorraine.fr)

Research laboratory: LCOMS (<https://lcoms.univ-lorraine.fr/>)

Description :

Title: Enhancing Multimodal Large Language Models through Grounding Techniques: A Comprehensive Evaluation and Optimization Framework

The proposed Master's thesis topic aims to advance the field of Multimodal Large Language Models (MLLM) by focusing on grounding techniques and their impact on model performance. The project will build upon the recent development of the MLLM Evaluation benchmark (MME), which addresses the need for a comprehensive evaluation of MLLMs' perception and cognition abilities.

One of the objectives of this project is to explore and implement various grounding techniques to enhance the perceptual and cognitive abilities of MLLMs. Grounding techniques involve linking linguistic elements to real-world entities, providing a more solid foundation for the model's understanding.

We should address data leakage concerns by manually designing instruction-answer pairs for the MME subtasks and developing a systematic approach to create concise and effective annotations, ensuring a fair comparison between different MLLMs without the need for extensive prompt engineering. We aim also to explore patterns and identify areas for improvement in existing models, providing insights into the potential directions for subsequent model optimization.

<https://arxiv.org/pdf/2306.13549.pdf>

<https://github.com/BradyFU/Awesome-Multimodal-Large-Language-Models/tree/Evaluation>

<https://cloud.google.com/blog/products/ai-machine-learning/how-to-use-grounding-for-your-llms-with-text-embeddings?hl=en>

Keywords : MLLM, Grounding techniques, Python.

Internship conditions:

Internship location and employer: LCOMS - IUT de Saint-Dié, 88100 Saint-Dié-des-Vosges, France

Compensation: Legal allowance.

Expected profile: Motivated master's/engineering student, volunteer, with skills in Python programming. LLM knowledge would be a plus.