

Rapport d'analyse de document PDF

Fichier: **Projet_NLP_1768509094377.pdf**

Type détecté: **article_scientifique**

Pages: **14**

Date d'analyse: **2026-01-15 22:32**

Résumé exécutif

Les modèles de langage de grande taille (LLMs) ont révolutionné le traitement automatique du langage naturel (TALN), mais les dialectes arabes, notamment la darija marocaine, sont insuffisamment représentés. Ce projet de recherche expérimentale vise à concevoir un mini-chatbot de question-réponse en darija marocaine, en utilisant des techniques de fine-tuning paramètre-efficace (LoRA et QLoRA) pour adapter le modèle ArabianGPT-0.3B-QA. L'objectif est de pallier le manque de ressources linguistiques structurées et les contraintes computationnelles. Les résultats montrent que LoRA offre des performances légèrement supérieures, tandis que QLoRA réduit significativement la consommation mémoire. L'interface Gradio démontre la cohérence et la compréhension des réponses du chatbot. Ce projet souligne la faisabilité et l'intérêt de développer des applications de TALN en darija marocaine avec des ressources limitées.

Points clés

- Les modèles de langage de grande taille (LLMs) ont transformé le TALN, mais les dialectes arabes comme la darija marocaine sont insuffisamment représentés.
- Objectif : concevoir un mini-chatbot de question-réponse en darija marocaine en utilisant des techniques de fine-tuning paramètre-efficace (LoRA et QLoRA).
- Approche : utilisation du modèle ArabianGPT-0.3B-QA avec des techniques de fine-tuning paramètre-efficace pour adapter le modèle à une tâche de question-réponse.
- Résultats : LoRA offre des performances légèrement supérieures, tandis que QLoRA réduit significativement la consommation mémoire.
- L'interface Gradio démontre la cohérence et la compréhension des réponses du chatbot en darija marocaine.
- Les méthodes de fine-tuning paramètre-efficace permettent d'adapter efficacement un modèle de langage arabe avec des ressources computationnelles limitées.
- Le projet démontre la faisabilité et l'intérêt de développer des applications de TALN en darija marocaine.

Informations extraites

Problème	Le manque de ressources linguistiques structurées en darija marocaine pour les systèmes de TALN.
Objectifs	Concevoir un mini-chatbot de question-réponse en darija marocaine, appliquer des techniques de fine-tuning paramètre-efficace.
Méthodes	-
Résultats	Le modèle affiné avec LoRA a montré des performances légèrement supérieures en termes de précision et de cohérence.
Conclusion	-
Mots-clés	darija, chatbot, question-réponse, ArabianGPT, LoRA, QLoRA, fine-tuning, PEFT, NLP, IA GPT

Annexe: Références de pages (approx.)

- Les modèles de langage de grande taille (LLMs) ont transformé le TALN, mais les dialectes arabes comme la darija marocaine sont insuffisamment représentés. (pages: 4, 14, support: fort)

- Objectif : concevoir un mini-chatbot de question-réponse en darija marocaine en utilisant des techniques de fine-tuning paramètre-efficace (LoRA et QLoRA). (pages: 4, 3, support: moyen)
- Approche : utilisation du modèle ArabianGPT-0.3B-QA avec des techniques de fine-tuning paramètre-efficace pour adapter le modèle à une tâche de question-réponse. (pages: 9, 4, support: moyen)
- Résultats : LoRA offre des performances légèrement supérieures, tandis que QLoRA réduit significativement la consommation mémoire. (pages: 11, 12, support: fort)
- L'interface Gradio démontre la cohérence et la compréhension des réponses du chatbot en darija marocaine. (pages: 12, 14, support: fort)
- Les méthodes de fine-tuning paramètre-efficace permettent d'adapter efficacement un modèle de langage arabe avec des ressources computationnelles limitées. (pages: 4, 5, support: moyen)
- Le projet démontre la faisabilité et l'intérêt de développer des applications de TALN en darija marocaine. (pages: 14, 6, support: moyen)