Abstract

- LLMs are a foundation for many NLP tasks but do not perform equally across dialects
- Dialect robustness across all tasks is key for inclusion of English speakers globally
- We release task agnostic dialect adapters which offer plug-and-play dialect robustness for anyone to use

Method

TADA creates modular adapters for task-agnostic dialect robustness in English NLP. We first generate a synthetic sentence-parallel corpus between dialects of English. We then align non-SAE dialects to SAE for task-agnostic transfer learning from SAE data.

Evaluations

- TADA can be mixed with Standard American English Task Adapters for any task
- We train and release adapters which improve performance for 4 major global dialects of English spoken by over 250 Million people across the globe.

Multi-Dialectal Evaluation

Comparison To Task-Specific Methods

Figure 1: An example of how negative inversion, a feature common in dialects of the US American South, can lead to incorrect classification

Figure 2: Average Performance across Multi-VALUE transformed GLUE tasks for 4 global Englishes

Figure 3: Comparison of TADA to Task-Specific dialect adaptation techniques on the 3 largest GLUE subtasks