Title: A Brief Study on Large Language Models for Text Classification

Introduction
Large Language Models (LLMs) like GPT and LLaMA have shown strong generalization in NLP tasks. This work briefly evaluates their zero-shot ability on text classification.

Method
We used two models: GPT-Neo and LLaMA-2. Given a sentence and label list, the models were prompted to choose the most fitting label without fine-tuning.

Experiment
On three datasets (AG News, Yelp, DBPedia), LLaMA-2 outperformed GPT-Neo. For example, LLaMA-2 achieved 75.1% accuracy on AG News, while GPT-Neo reached 68.4%.

Conclusion
LLMs can classify text reasonably well in zero-shot settings. Larger models perform better, but even smaller ones show promise.