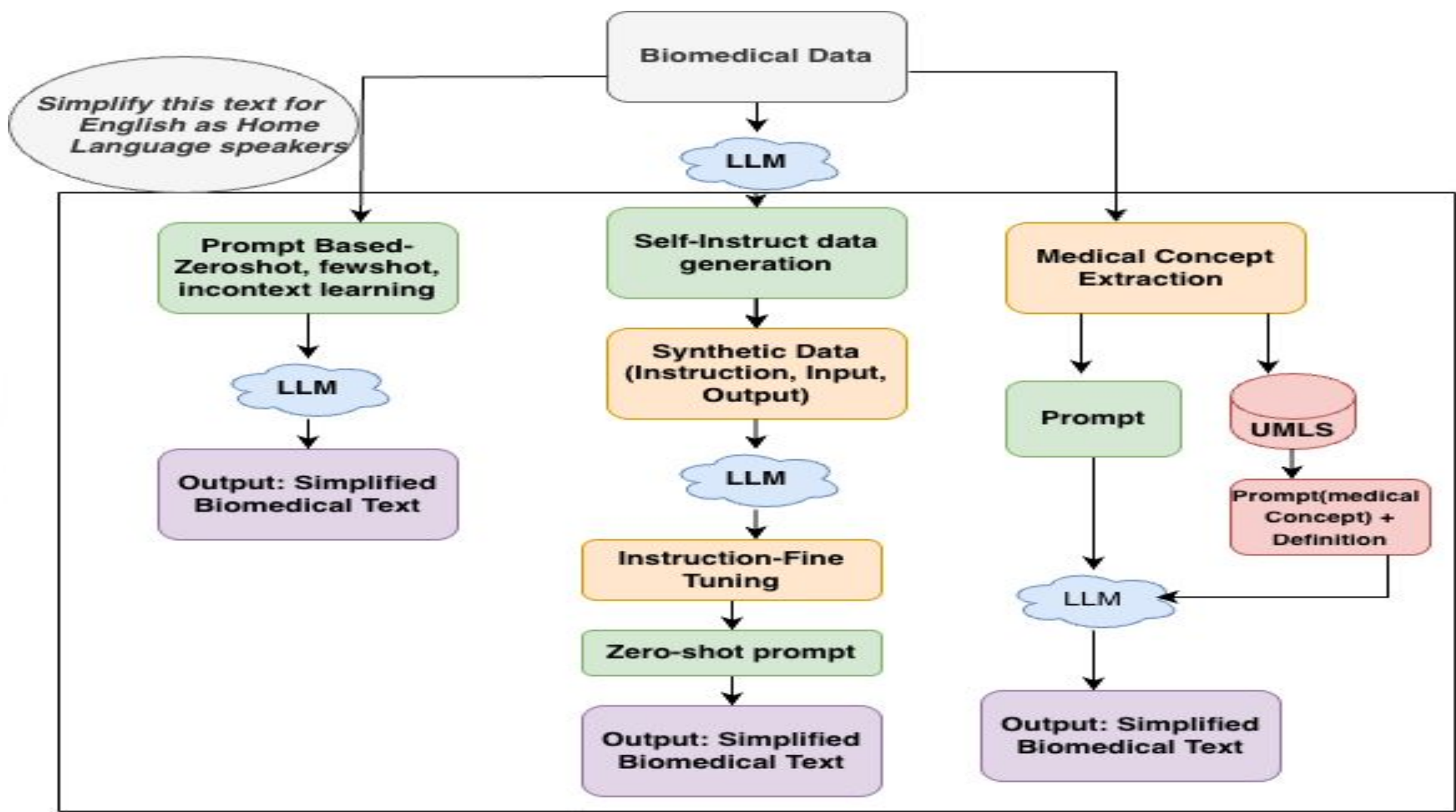


Evaluating Large Language Models for Multilevel Biomedical Text Simplification

Student: Nomonde Khalo
Supervisor: Dr Jan Buys



Biomedical texts are complex and inaccessible to non-experts. Large Language Models (LLMs) show promise for text simplification, but often fail to adapt to diverse reader proficiencies.



Research Question: To what extent does fine-tuning large language models improve readability and comprehension of biomedical texts for users with diverse Home Language and First Additional Language proficiencies?

Methodology

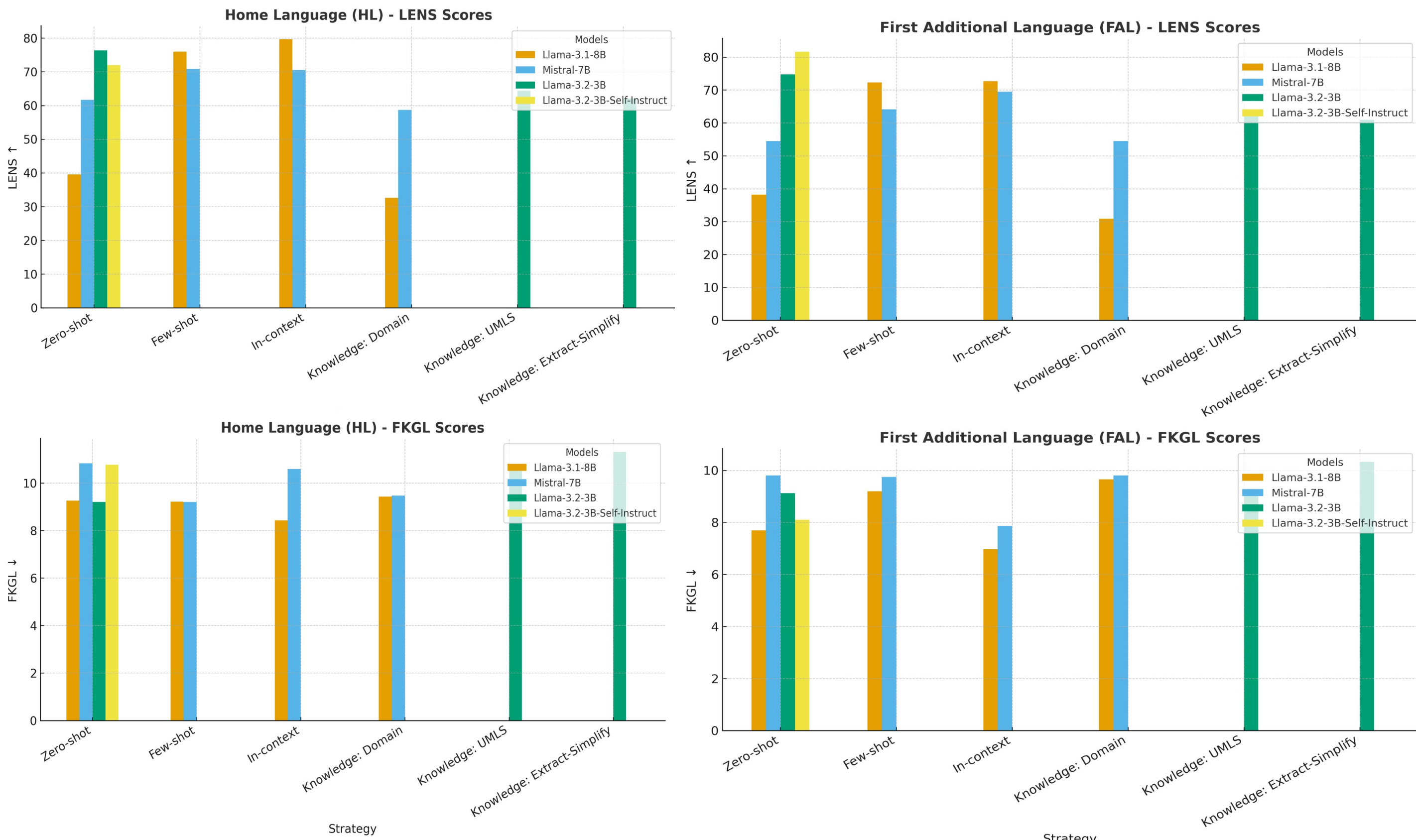
Proficiency-Aware

Simplification: We analyze prompt-based strategies for tailoring medical text to English Home Language (HL) and First Additional Language (FAL) readers, balancing readability and content fidelity.

Efficient Model Tuning: A small Self-Instruct-tuned model achieves comparable readability for FAL users, outperforming larger general and domain-specific models.

Domain Knowledge Impact: We evaluate how biomedical knowledge—via domain-specific LLMs and external knowledge integration—affects simplification quality and reliability.

Results



HL Self-Instruct Simplified Output

49 studies with 12,067 patients were reviewed. 38 studies compared ACE inhibitors (ACEi) with a placebo, 4 compared angiotensin II receptor antagonists (AIIRA) with a placebo, and 7 compared ACEi and AIIRA directly.

FAL Self-Instruct Output

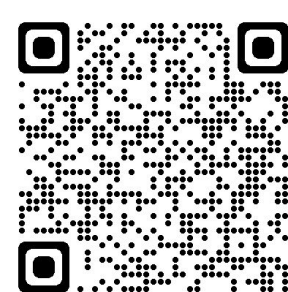
There were 49 studies about patients with high blood pressure and kidney problems. 38 studies compared the use of ACE inhibitors (medications that help the heart and blood vessels relax) with a placebo (a dummy treatment)

Conclusion

- General-purpose models like Mistral excel at preserving semantic content, while in-context prompting of models like Llama-3.1-8B achieves the highest readability scores.
- Conversely, a small, self-instruct tuned model (Llama-3.2-3B) achieves comparable readability for FAL audiences.

Acknowledgement

This work is based on research supported by the Hasso Plattner Institute for Digital Engineering, through the HPI Research School at UCT



Authors: Nomonde Khalo, KHLNOM004@myuct.ac.za
Dr Jan Buys, jbuys@cs.uct.ac.za