Natural Language Processing Frontiers in African Indigenous Languages of Uganda

Nakirizi Nalwoga[1,2]**, Otim Oryondu**[3]**, Kizza Kibibi**[2,3]

[1] Makerere University, Kampala

[2] Uganda Christian University, Mukono

[3] Makerere University Business School (MUBS)

Author notes

Nakirizi Nalwoga is affiliated with Makerere University, Kampala and focuses on Computer Science research in Africa.
Otim Oryondu is affiliated with Makerere University Business School (MUBS) and focuses on Computer Science research in Africa.
Kizza Kibibi is affiliated with Uganda Christian University, Mukono and focuses on Computer Science research in Africa.

Abstract

Natural Language Processing (NLP) has shown promise in transforming how information is processed and utilised across various languages globally. However, its application in African indigenous languages remains underexplored, particularly in Uganda where multilingualism is prevalent. A comprehensive literature review was conducted to identify existing research on NLP applications in Ugandan indigenous languages. A series of interviews with language experts and software developers were also carried out to gather insights into the current landscape and future prospects. While significant progress has been made, there is a notable lack of standardised datasets for training models specific to these languages, which limits model performance and generalisation across different dialects. The findings suggest that while challenges remain, concerted efforts towards developing robust NLP solutions tailored to Ugandan indigenous languages could lead to substantial advancements in multilingual information processing. Developing a collaborative framework between academia and industry stakeholders is recommended to accelerate the creation of reliable datasets. Furthermore, promoting research funding specifically for African languages within the context of NLP can be beneficial. Model estimation used $\hat{\{\theta\}} = argmin\{\theta\} sumiell(yi, f\theta(\xi)) + \lambda lVert\theta rVert\,2^2$, with performance evaluated using out-of-sample error.

Keywords: *African, Computational Linguistics, Morphology, Pragmatics, Syntax, Typology, Corpus Linguistics*

ABSTRACT-ONLY PUBLICATION

This is an abstract-only publication. The complete research paper with full methodology, results, discussion, and references is available upon request.

✉ REQUEST FULL PAPER

Email: info@parj.africa

Request your copy of the full paper today!

SUBMIT YOUR RESEARCH

Are you a researcher in Africa? We welcome your submissions!

Join our community of African scholars and share your groundbreaking work.

Submit at: app.parj.africa

Scan to visit app.parj.africa

Open Access Scholarship from PARJ

Empowering African Research | Advancing Global Knowledge