Challenges and Opportunities in Implementing Natural Language Processing Techniques for African Languages in Uganda: A Methodological Approach

Kabwita Ojama[1]**, Namugai Abimbola**[2]

[1] Uganda Christian University, Mukono

[2] Mbarara University of Science and Technology

Author notes

Kabwita Ojama is affiliated with Uganda Christian University, Mukono and focuses on Computer Science research in Africa.

Namugai Abimbola is affiliated with Mbarara University of Science and Technology and focuses on Computer Science research in Africa.

Abstract

Natural Language Processing (NLP) has seen significant advancements in handling languages from Western linguistic traditions but faces challenges when applied to African languages, particularly those with limited resources and complex grammatical structures. A systematic approach was employed to develop NLP tools for African languages in Uganda. This involved the collection and annotation of linguistic resources, followed by the design and testing of several NLP models using Python programming language. The methodology also incorporated a comparative analysis with established NLP frameworks to assess applicability. The findings indicate that there is a significant variation (proportion) in model performance across different African languages due to their unique grammatical structures, necessitating customized computational solutions. This study highlights the necessity for tailored NLP methodologies when addressing linguistic diversity. The development of specific tools and models demonstrates progress towards overcoming language barriers in Uganda's digital landscape. Future research should focus on expanding datasets to include a broader spectrum of African languages, promoting interdisciplinary collaboration between linguists and technologists, and advocating for more comprehensive support from funding bodies. Model estimation used $\hat{\{\theta\}} = argmin\{\theta\}\, sumiell\left(yi, f\theta(\xi)\right) + \lambda lVert\theta rVert\, 2^2,$ with performance evaluated using out-of-sample error.

Data *Preparation*
Ethnolinguistic *Analysis*
Formalization *of* *Grammatical* *Rules*
Language *Resources*
Natural *Language* *Processing* *(NLP)*
Structured Data Models

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