UGANDA CHRISTIAN UNIVERSITY

FACULTY OF ENGINEERING, DESIGN AND TECHNOLOGY

Program: Bachelor of Science in Computer Science
Course: Big Data Mining and Analytics

Building natural language processing tools for Runyakitara

Name: Yawe Arthur Shalom
Reg Number: M23B23/016

Short essay describing the insights you derived

The review article provides deep insights into the intersection of technology, linguistics, and cultural preservation. It demonstrates how Natural Language Processing (NLP) can play a vital role in revolutionizing indigenous languages that are under threat of extinction due to globalization and limited educational use. The authors present a practical and inspiring case for how computational tools can empower both language learners and native speakers, while preserving cultural identity.

One key insight that has been presented is the importance of developing digital linguistic resources for under looked languages such as Runyakitara. Unlike major world languages like English, the Runyakitara language group lacks digitized texts, corpora, and computational tools. The authors emphasize that before any advanced NLP systems can be developed, it is essential to build foundational resources like corpora and morphological analyzers. This process forms the backbone for developing applications such as spell-checkers and grammar-checkers. From this, I learned that successful language technology depends on data without a properly constructed corpus, even the most advanced tools cannot function effectively.

Another important insight is the vital role of technology in language preservation and revitalization. The article highlights that many African languages, though spoken by millions, face endangerment because they are not widely taught or used in formal communication. Through the creation of tools like RunyaSpeller and LearnRunya, the authors illustrate how technology can support both literacy and language learning. I realized that digital tools not only make learning more accessible but also enhance the prestige and visibility of indigenous languages, encouraging younger generations to embrace them.

Additionally, the article provides a valuable understanding of the social and educational challenges facing indigenous languages in Uganda. Despite government policies promoting "mother tongue" education, the dominance of English and Kiswahili continues to marginalize languages like Runyakitara. This has led to reduced intergenerational transmission and a growing preference for foreign languages. The authors' approach shows that language revitalization must go beyond policy it requires innovative educational technology, community involvement, and cultural motivation.

Finally, I found the integration of linguistics and artificial intelligence particularly inspiring. The development of the RunyaMorph analyzer and the use of NLP to process morphologically rich Bantu languages reveal how modern AI techniques can adapt to diverse linguistic structures. It shows that language technology is not limited to Western languages but can be localized to serve unique African linguistic systems. This insight expands the vision of what AI can achieve in preserving global linguistic diversity.

In conclusion, the article not only deepened my appreciation for the role of computational linguistics in language preservation but also reinforced the idea that technology, when thoughtfully applied, can protect cultural heritage and promote inclusive education. The project stands as a model for how linguistic research and innovation can converge to support both academic and social progress in Africa.

Recommendations to the Ethnic Leaders of the Runyakitara Speakers on the Use of NLP Models for Language Preservation

1. Support the Creation and Expansion of Runyakitara Digital Resources
Ethnic leaders should mobilize their communities to contribute to building digital corpora by collecting oral histories, songs, folktales, and traditional stories in Runyakitara. These materials can be digitized and integrated into the *RunyaCorpus*, providing valuable data for NLP tools such as morphological analyzers, spell-checkers, and grammar systems.

2. Promote the Integration of NLP Tools in Education
Leaders should advocate for the inclusion of tools like *RunyaSpeller* and *LearnRunya* in schools to enhance language learning and literacy. By encouraging teachers and learners to use these tools, they can strengthen the role of Runyakitara as a medium of education and cultural identity.

3. Encourage the Use of Technology in Media and Public Communication
The use of NLP models can extend to digital media platforms, radio, and newspapers. Ethnic leaders should promote the use of Runyakitara in online publications and local broadcasts through speech recognition, text generation, and translation tools to make the language more visible in public life.

4. Partner with Researchers and Developers
Collaborations with universities, linguists, and technology developers can help expand ongoing projects such as *RunyaMorph* and the *RunyaCorpus*. Such partnerships ensure technical accuracy, sustainable progress, and opportunities for funding and innovation in Runyakitara NLP development.

5. Promote Community Awareness and Cultural Pride
Language preservation depends on community participation. Leaders should organize cultural events and campaigns that emphasize the value of speaking and writing in Runyakitara. Encouraging the use of NLP tools on mobile devices can help younger generations see their language as modern and relevant.

6. Advocate for Government and Institutional Support
Ethnic leaders should work with government bodies and educational institutions to secure policy support and funding for indigenous language digitization. Promoting the inclusion of NLP-based tools in national education policies can help ensure long-term sustainability for language revitalization.

Conclusion
Ethnic leaders of the Runyakitara-speaking communities play a vital role in linking cultural heritage with modern technology. By supporting NLP initiatives, they can help preserve their language for future generations while enhancing literacy, education, and cultural pride through digital innovation.