
Contribution ID: **130** Type: **Talk**

# Transformer-based Sense Embeddings with Deep Learning Large Language Models for Low-Resource Language Disambiguation

*Wednesday, 4 December 2024 16:30 (20 minutes)*

Determining a word's accurate meaning in each context is known as Word Sense Disambiguation (WSD)[1], and it is one of the most significant problems in Natural Language Processing (NLP)[2]. This undertaking is particularly challenging for low-resource languages like Sesotho sa Leboa since there are few annotated corpora and linguistic resources available for them. This study explores the application of many transformer-based and deep learning models for WSD in Sesotho sa Leboa, with good results despite the language's resource constraints. This study employs a variety of deep learning architectures, including transformer-based models such as Recurrent Neural Networks with Long Short-Term Memory (RNN-LSTM), Bidirectional Gated Recurrent Units (BiGRU), and an LSTM-based Language Model (LSTMLM), as well as models like DistilBERT with Naive Bayes (DistilBERT & NB), DeBERTa, T5, and ALBERT[3][4].

The study makes use of the unique hardware characteristics of the T4 GPU to improve and optimize the runtime of deep learning language models, especially big transformers. The purpose of the NVIDIA T4 Tensor Core GPU is to speed up deep learning and machine learning operations. It works especially well for training and inferring big language models. Every phase entail making efficient use of software optimizations in addition to comprehending and utilizing hardware features. The BiGRU model outperformed other deep learning language models with an accuracy of 79%, demonstrating the effectiveness of bidirectional processing effectively capturing contextual information. With an accuracy of 70%, DeBERTa beat the other transformer-based large language models to enhance pre-training techniques that prioritize spatial and contextual embeddings.

**Keywords:** Word Sense Disambiguation, Sesotho sa Leboa, Low-Resourced Languages, Deep Learning, Transformer Models, RNN-LSTM, BiGRU, DeBERTa, NLP

## Student or Postdoc?

## Email address

## Co-Authors

## CHPC User

## CHPC Research Programme

## Workshop Duration

**Primary authors:** MASETHE*, Hlaudi (Tshwane University of Technology); Mr MASETHE, Hlaudi Daniel (Tshwane University of Technology)

**Presenter:** MASETHE*, Hlaudi (Tshwane University of Technology)

**Session Classification:** DIRISA

**Track Classification:** DIRISA