**Using NLP on multimodal data analysis enhances sentiment recognition accuracy of Afrikaans music videos, showing that analysing various media formats provides deeper insights and improves personalised music recommendations.**

# Multimodal sentiment analysis of Afrikaans musical videos based on personal preferences.

## INTRO

- Our project addresses gaps in music video recommendation systems for low-resourced languages like Afrikaans.
- Current systems often fail to align with users' unique preferences, leading to suboptimal experiences.
- There is a need for quick, personalized recommendations to give users better control over their content.

## METHODS

- Audio and visual data were extracted from the 117 segmented YouTube music videos, along with song lyrics, ensuring diverse sentiment representation.
- Employed sentiment analysis techniques using NLP, including Multinomial Naive Bayes and KNN classifiers for lyrics, Keras model for audio, and VGG16 model for images.
- Averaged sentiment results from text, audio, and image analyses to refine song recommendations based on user preferences.

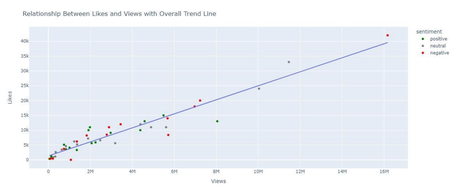


Figure 1: Relationship between likes and views with trend line

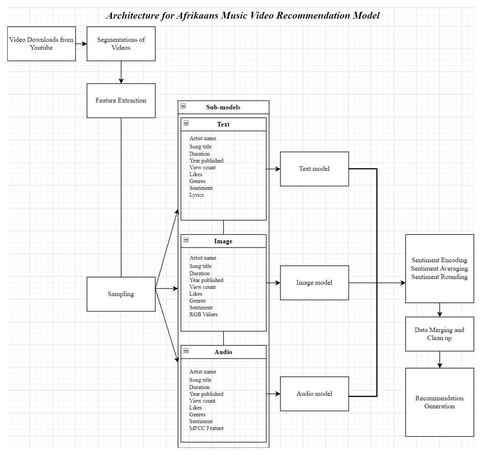


Figure 2: Architecture for Afrikaans music video recommendation model.

## RESULTS

| S/N | Artist Name | Song Title | Duration | Year Published | View count | Likes | Genres | Text sentiment prediction | Audio sentiment prediction | Overall sentiment | Sentiment label |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Juan Boucher | Whiskey en jy | 00:04 | 2019 | 5,714,448 | 8400 | Pop Rock | 0.0 | 1.6 | 1.0 | Neutral |
| 2 | Bernice West | As hy weer kom | 00:03 | 2023 | 1,912,216 | 10000 | Gospel | 2.0 | 1.4 | 2.0 | Positive |
| 3 | Caroline Grace | Voor jou | 00:03 | 2017 | 635,483 | 3400 | Pop | 1.0 | 0.8 | 1.0 | Neutral |

Figure 3: Over-all sentiment prediction
Overall sentiment prediction from the text and audio model

## DISCUSSION

- Our recommendation system includes two variations: Variation 1 allows users to input preferred sentiment, genre, and date to receive tailored song recommendations. Variation 2 accepts a YouTube URL, provides metadata (artist, song title, year of upload, predicted sentiment), and recommends similar songs.

## VUSUALIZATIONS


Figure 4: Word cloud depicting all native speaker annotated lyrics on sampled data

```
Enter the YouTube URL of the song: https://www.youtube.com/watch?v=cUmald-n7Ic&list=PL9Zt13PbXnLeckQc6n9zcTCL-dDUjpoAg
Extracted - Title: Madelein (Official Music Video), Artist: Brendan Peyper, Genre: pop, Sentiment: positive, Year: 2024

Recommendations based on your song:
17 shots by Christia Visser - Genre: pop
as hy weer kom by Bernice West - Genre: gospel
asemrouend by Leon Ferreira - Genre: gospel
beter helfte (amptelike musiek video) by Bobby van Jaarsveld - Genre: pop
bietjie meer van jou by 4Werke - Genre: pop
daal ding by Ivan Roux - Genre: dance
die eerste keer by Arno Jordaan - Genre: pop
dis liefde by Ivan Roux - Genre: pop rock
droë riviere by Liezel Pieters - Genre: pop
een vir een by Bobby van Jaarsveld - Genre: pop
ek bid vir jou by Caroline Grace - Genre: pop
ek kan nie ophou kyk (official music video) by Franja du Plessis - Genre: pop
```

Figure 5: Afrikaans music recommendation system variation 2
Application of YouTube URL to provide metadata like artist, song title, and upload year, predicts song sentiment, and suggests similar songs.

```
Available Sentiments:  ['negative' 'positive' 'neutral']
Select your preferred sentiment: positve
Available Genres:  ['pop' 'gospel' 'dance' 'rock' 'pop rock' 'folk' 'alternative']
Select your preferred genre: pop
Available Years:  [2019 2023 2015 2016 2014 2018 2017 2022 2010 1994 2020 2011 2013]
Select your preferred year of publication:
```

Figure 6: Afrikaans music recommendation system variation 1
Music recommendations are generated by filtering a data frame according to the user's input choices, and results are displayed within Google Colab.

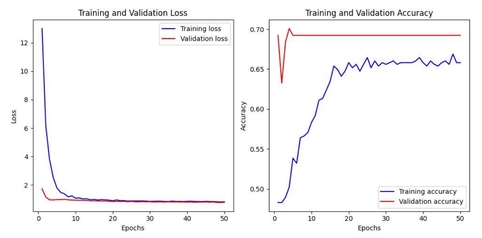


Figure 7: Loss and Accuracy plot for the Keras audio model
Figure 7 displays the training and validation loss and accuracy of the Keras audio model over 50 epochs, highlighting the model's rapid improvement in learning and its ability to generalize well across the dataset.

**Refiloe Kekana, Abiola Yetunde Akinbowale, Pholosho Masilane**

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Department of Computer Science

Faculty of Engineering, Built Environment and Information Technology

Fakulteit Ingenieurswese, Bou-omgewing en Inligtingtegnologie / Lefapha la Boetšenere, Tikologo ya Kago le Theknolotši ya Tshedimošo

Capstone Project - MIT 808

Course Coordinators:
Dr. Vukosi Marivate (vukosi.marivate@cs.up.ac.za)
Abiodun Modupe (abiodun.modupe@cs.up.ac.za)

Scan me