

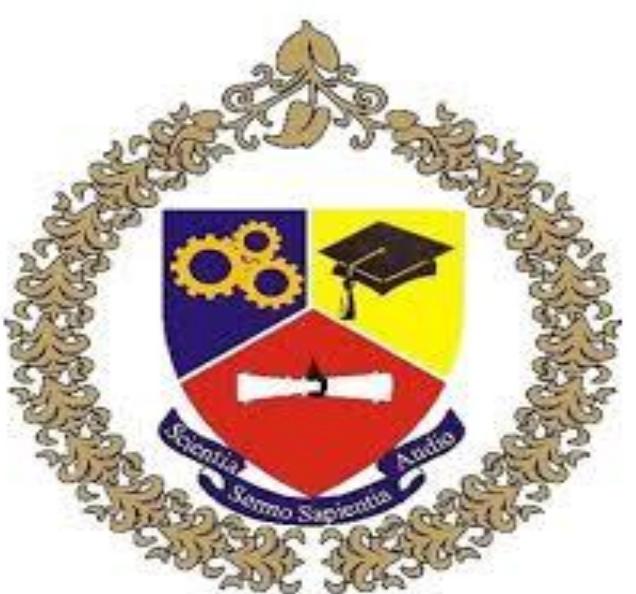


Green Topics, Deep Roots: Energy-Aware Topic Modelling of Multilingual Nigerian Lyrics

Sakinat Oluwabukonla Folorunso¹, Tosin Sina Akerele¹, Francisca Onaolapo Oladipo², Oluwakemi Rukayat Giwa¹

¹Olabisi Onabanjo University, Nigeria; ²Thomas Adewumi University, Oko-Irese, Nigeria

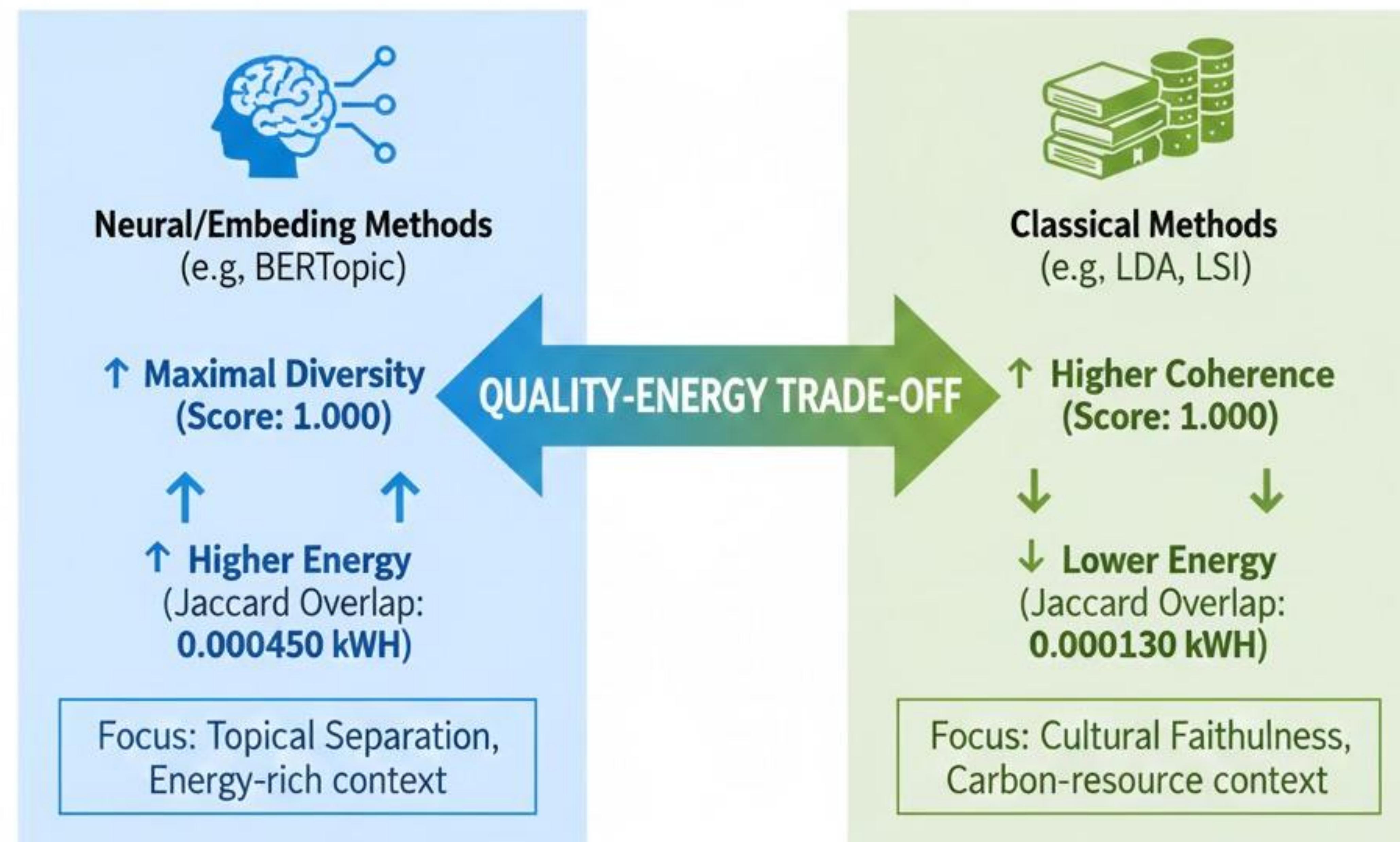
(sakinat.folorunso, tosin.akerele)@ouagoiwoye.edu.ng, francisca.oladipo@tau.edu.ng, oluwakemiemida@gmail.com



GRAPHICAL ABSTRACT



TOPIC MODELING NIGERIAN LYRICS: THE QUALITY-ENERGY TRADE-OFF



Study analyzes themes in Nigerian lyrics (English, Pidgin)
various topic models.

Practical guidance for sustainable, human-centred cultural metering



Green Topics, Deep Roots: Energy-Aware Topic Modelling of Multilingual Nigerian Lyrics

Sakinat Oluwabukonla Folorunso¹ Tosin Sina Akerele¹ Francisca Onaolapo Oladipo² Oluwakemi Rukayat Giwa¹

¹Olabisi Onabanjo University, Nigeria ²Thomas Adewumi University, Oko-Irese, Nigeria

{sakinat.folorunso, tosin.akerele}@ouagoiwoye.edu.ng
francisca.oladipo@tau.edu.ng, oluwakemi@midia@gmail.com

Abstract

We investigate how to model themes in Nigerian lyrics while respecting energy limits faced in low-resource settings. Our multilingual corpus spans English, Yoruba, and Nigerian Pidgin, including everyday code-switches and devotional terms, to preserve cultural nuance. We benchmark seven topic models (NMF, LDA, LSI, HDP, BERTopic, Top2Vec, GSDMM). Methods combine standard semantic metrics—coherence (C_v , UMass), topic diversity, and Jaccard overlap—with direct energy measurements (kWh). Results show a pronounced quality–energy trade-off: NMF achieved the highest coherence among classical models ($C_v = 0.6045$) at $\sim 2 \times 10^{-6}$ kWh, while LSI was similarly frugal with competitive quality. By contrast, BERTopic delivered maximal diversity (1.000) with disjoint topics (Jaccard = 0.000) but at markedly higher energy (0.000450 kWh). Top2Vec underperformed on coherence ($C_v = 0.2698$) and consumed more energy than most classical baselines (0.000113 kWh); GSDMM drew the most energy (0.000509 kWh) with undefined coherence on this short, sparse corpus. Interpreting these findings, we argue that in contexts where electricity and computing are scarce, classical models—particularly NMF—offer a culturally faithful, carbon-conscious starting point, while neural or embedding-based methods may be reserved for cases that demand maximal topical separation. Our study offers practical guidance for teams seeking sustainable, human-centred text mining of indigenous cultural materials.

Keywords— Green AI; Energy-aware NLP; Topic Modelling; Nigerian Lyrics; Low-Resource Settings.



INTRODUCTION

 Creative AI reveals culture

 Computing = limited electricity & compute

 Standard NLP pipelines often flatten multilingual nuances

 Which models best maps themes in a cost-efficient manner?

 Which models best maps themes in a cost-efficient manner?

Motivation

- a. The benefits of Creative AI are unevenly distributed, where electricity, hardware, and annotated data are scarce.
- b. This challenge is especially acute for multilingual Nigerian music, where lyrics braid Yoruba, English, and Nigerian Pidgin into rich code-switched expressions.
- c. Standard NLP pipelines often flatten or overlook these cultural and linguistic nuances.
- d. There is a need for energy-conscious, culturally aware AI approaches that support inclusion, equity, and heritage preservation.

Problem Statement

Which topic-modelling methods best surface culturally meaningful themes while staying affordable to run in low-resource environments?

Aim

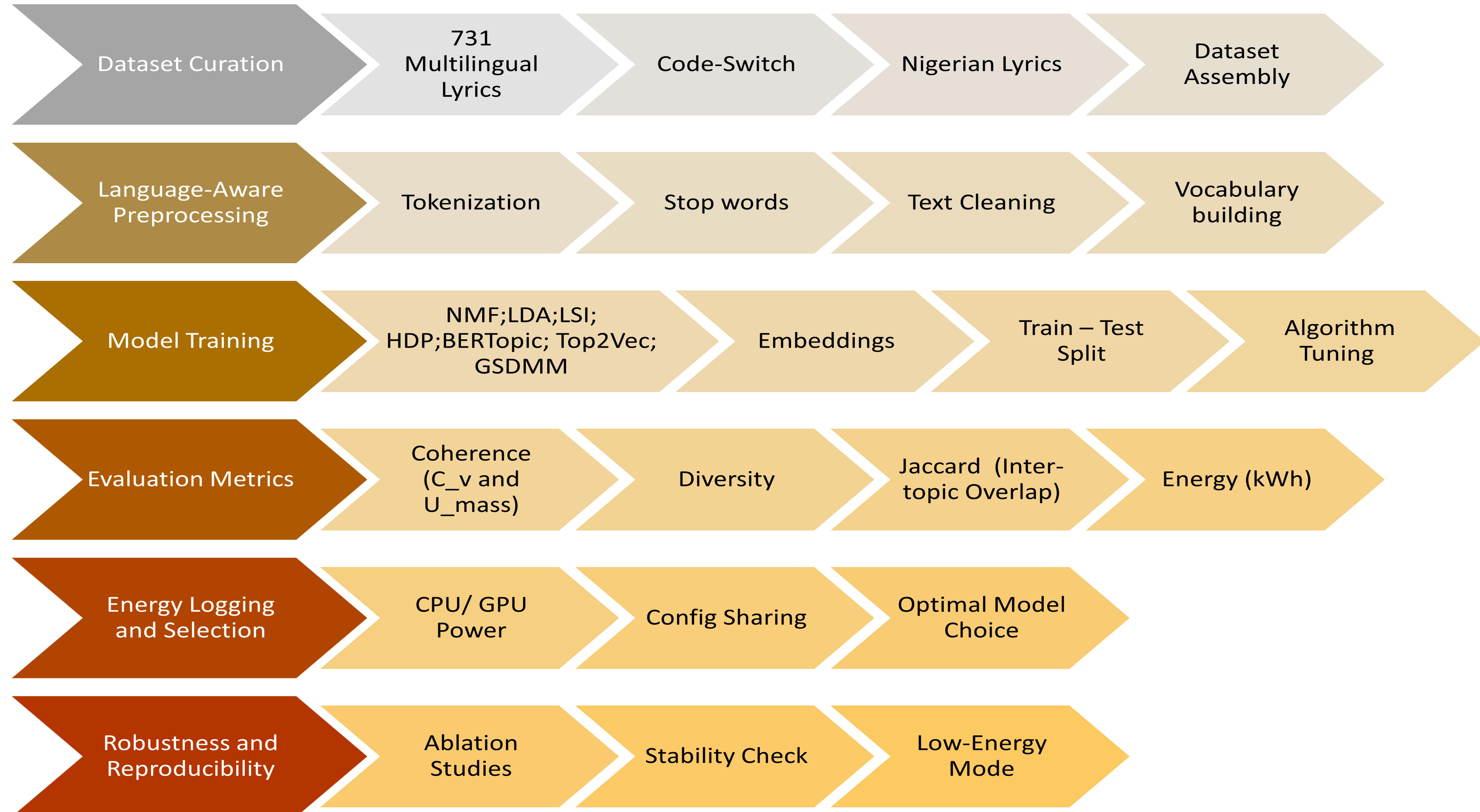
To identify energy-efficient and culturally faithful topic-modelling methods for multilingual Nigerian lyrics that are practical for use in low-resource environments.

Objectives

- a. Curate and preprocess a multilingual Nigerian lyrics corpus that preserves code-switching and cultural nuance.
- b. Benchmark seven topic-modelling approaches using both semantic and energy metrics.
- c. Log and compare energy consumption across models.
- d. Evaluate coherence, diversity, and overlap to examine quality-energy trade-offs.
- e. Recommend sustainable NLP workflows for cultural preservation and community applications.



ENERGY-AWARE TOPIC MODELLING PIPELINE





RESULTS

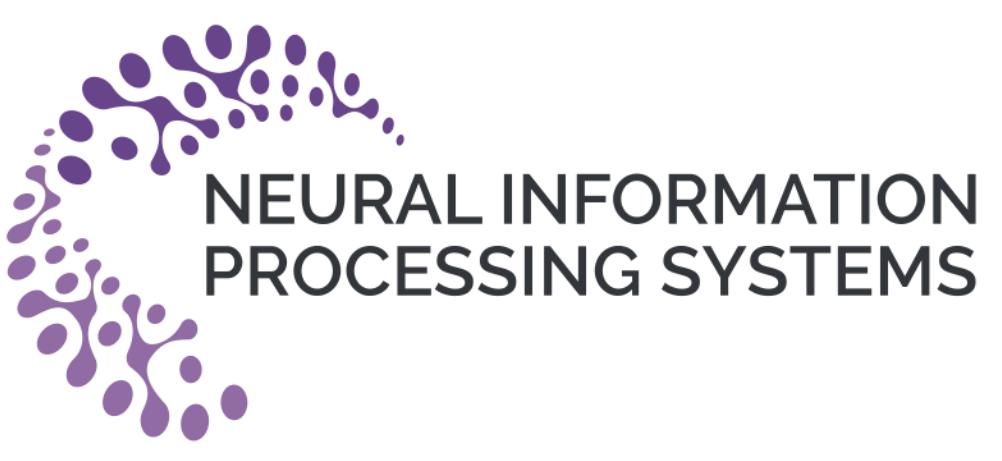


Table 1: Topic modelling results on Nigerian lyrics

Model	$C_v \uparrow$	$U_{\text{mass}} \uparrow$	Diversity \uparrow	Jaccard \downarrow	Energy (kWh)
NMF	0.604543	-2.851108	0.900	0.028173	0.000002
LDA	0.440793	-1.767695	0.600	0.213304	0.000047
LSI	0.487633	-1.392919	0.600	0.215767	0.000001
HDP	0.527935	-1.962347	0.360	0.502880	0.000013
BERTopic	0.578891	-1.834889	1.000	0.000000	0.000450
Top2Vec	0.269799	-6.364287	0.675	0.213282	0.000113
GSDMM	NaN	NaN	0.775	0.067355	0.000509

Higher is better for C_v , UMass (less negative), and Diversity; lower is better for Jaccard and Energy.

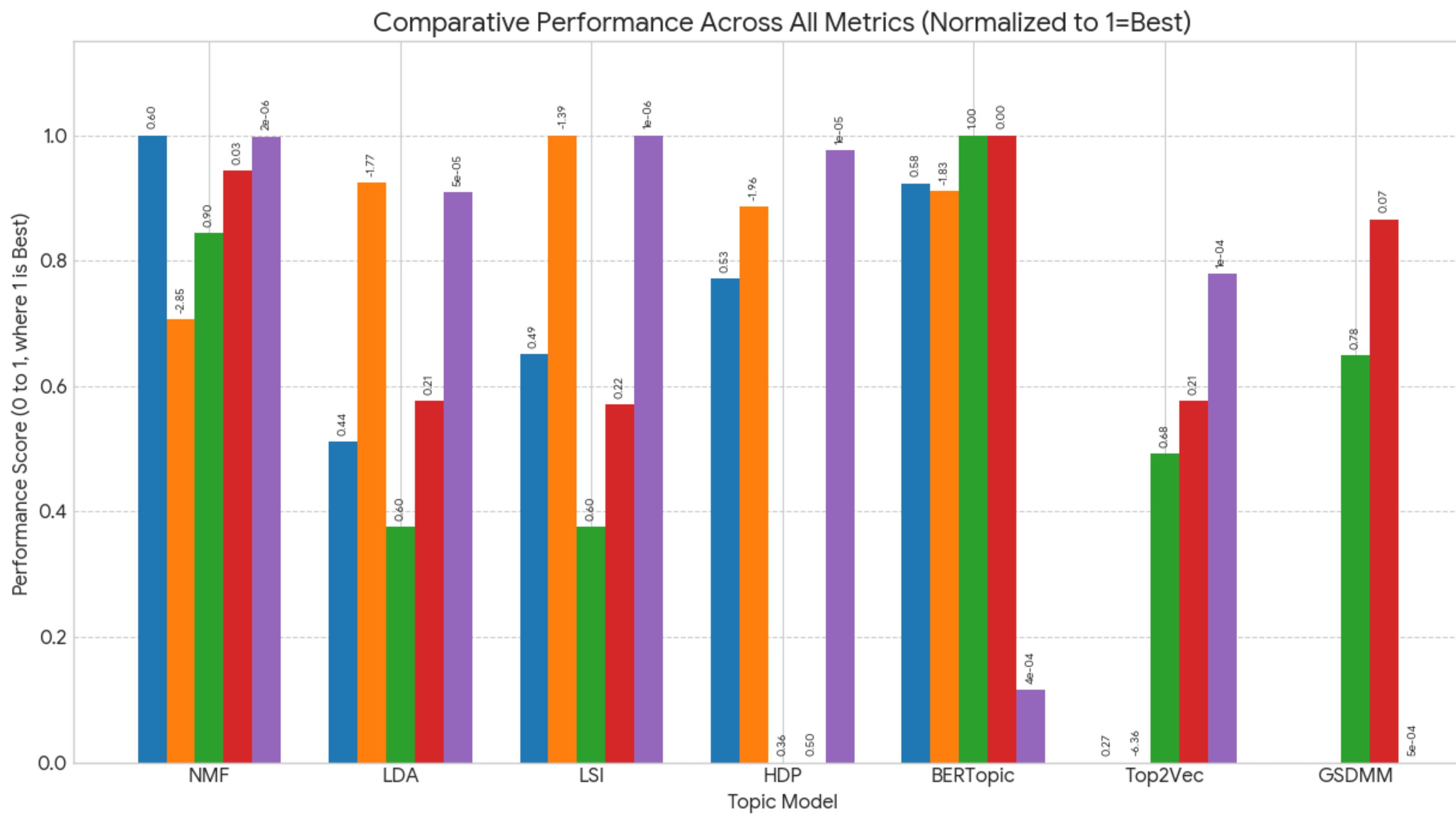


Figure 1: Word clouds across models.

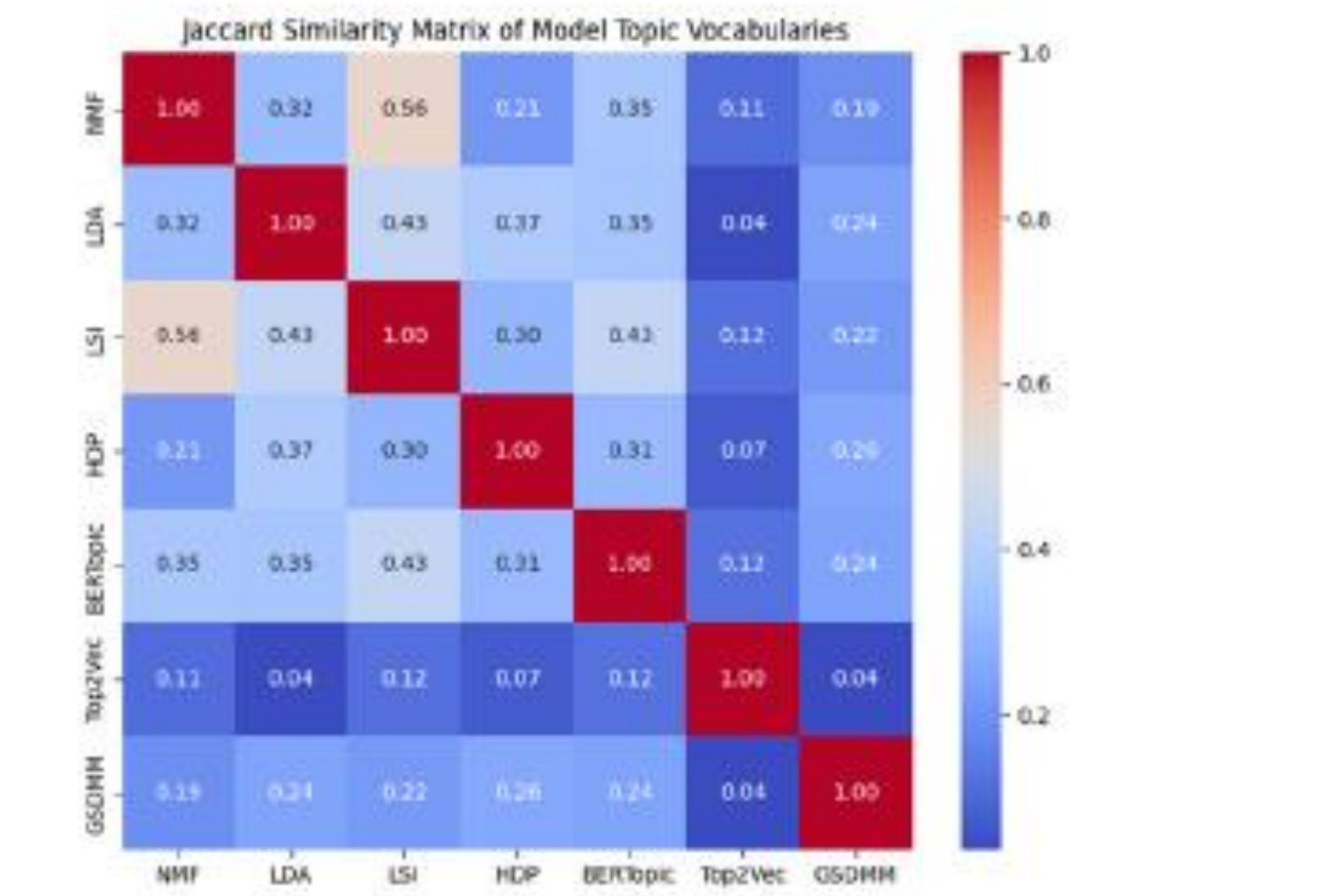
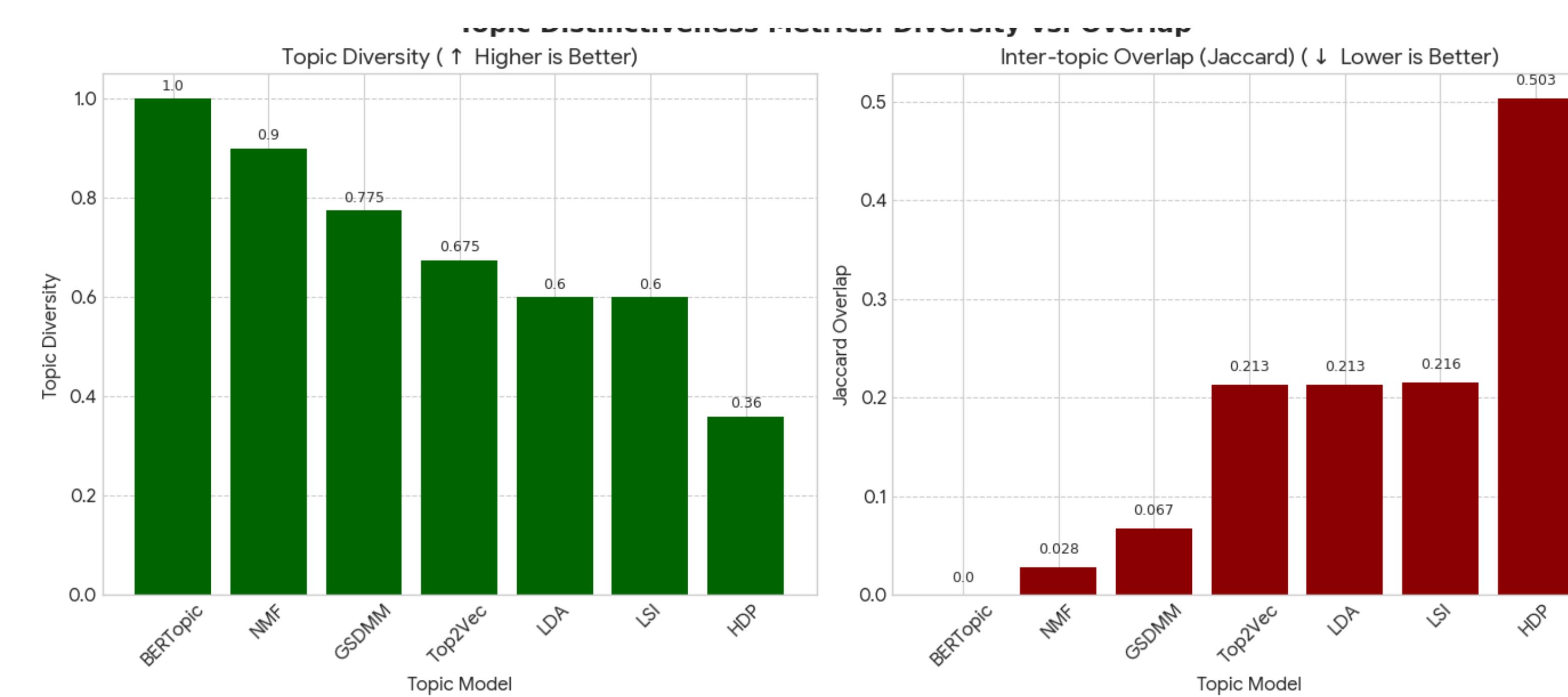


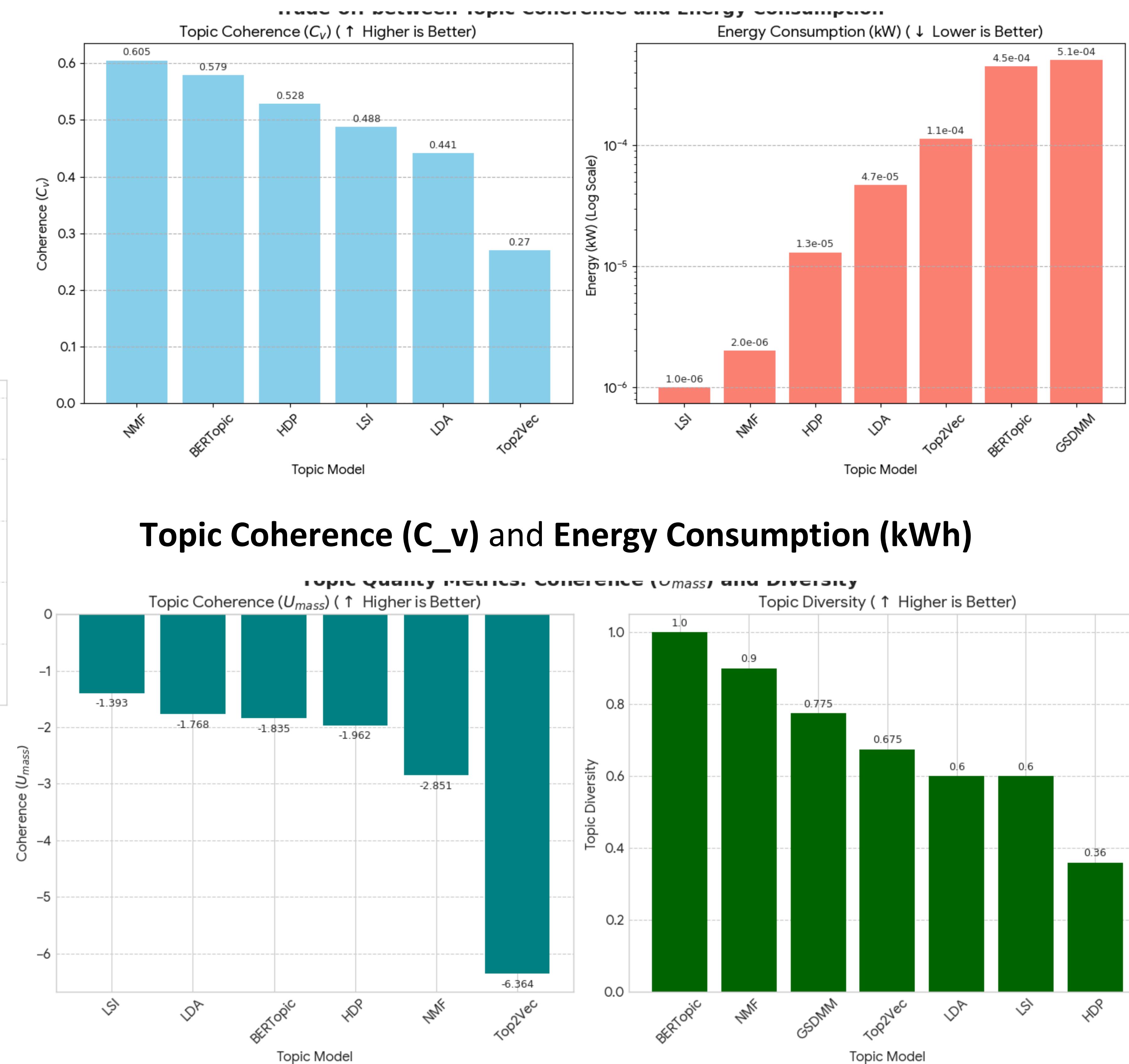
Figure 2: Jaccard similarity heatmap.



RESULTS



Topic Diversity and Inter-topic Overlap (Jaccard)



Coherence (U_{mass}) and Topic Diversity



Discussion

A clear hierarchy emerges for short, code-switched Nigerian lyrics when meaning and energy both matter

Embedding-first pipelines

- BERTopic produces crisp, disjoint clusters valuable for exploration but at higher energy cost
- Top2Vec struggles on very short, code-mixed inputs



Probabilistic models

- LDA stays frugal but topics blur
- HDP raises coherence but adds redundancy



Classical models

- NMF delivers the clearest, most coherent themes at milliwatt-hour scale

Why this matters:

- Lyric lines often weave English, Yoruba, and Nigerian Pidgin
- Heavy cross-Lingual models aren't strictly needed when language-aware preprocessing is strong
- Classical methods offer culturally faithful topics with tiny energy budgets

Modeling themes in Nigerian lyrics with energy constraints

Multilingual corpus

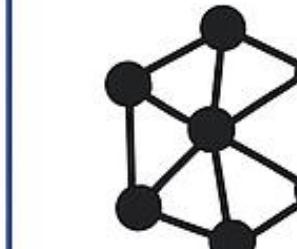


with everyday code-switching

Trade-offs

- Coherence (C_y)
- Topic diversity (\uparrow)
- Jaccard overlap
- Energy (kWh)

Topic modeling benchmark



NMF
LDA
LSI
HDPTopic
Top2Vec
GSDMM

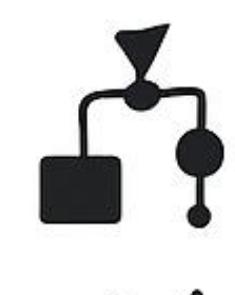


- Topic diversity (\uparrow)
- Jaccard overlap (\downarrow)

Recommendations



Classical models (esp. NMF) are carbon-conscious

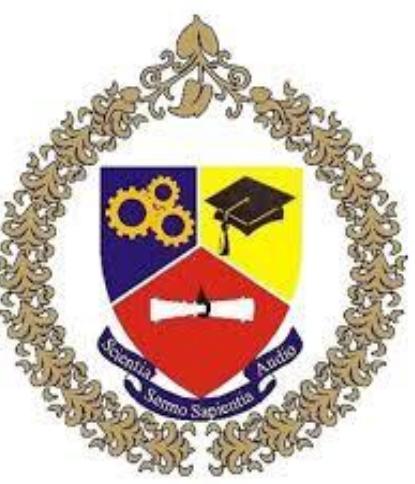


Character-redundancy (\uparrow)



Energy (kWh)

- Classical models (esp. NMF) are carbon-conscious
- Embedding-based models yield separation

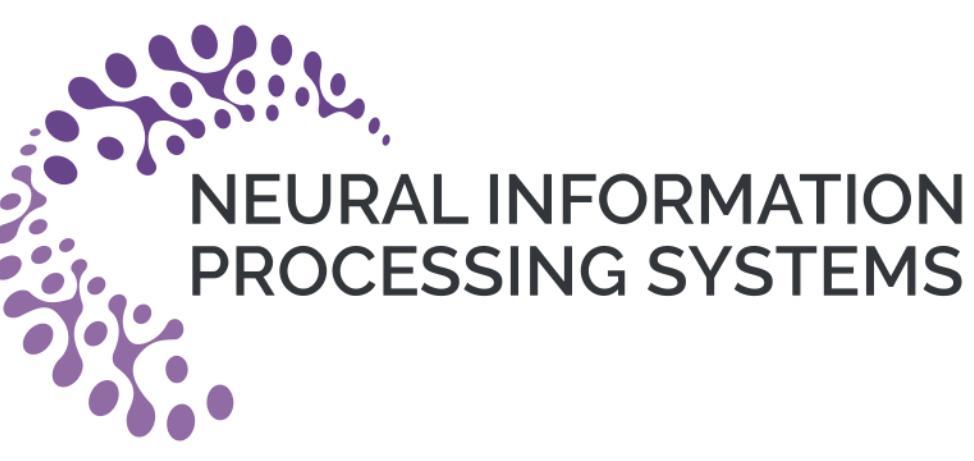


Comparison to Existing Work (with Metrics)

Relative to a classical LDA baseline on our corpus, the metric deltas align with expectations from prior work. NMF improves coherence by +0.164 Cv ($0.6045 \rightarrow 0.4408$), raises diversity by +0.300 ($0.90 \rightarrow 0.60$), and lowers lexical overlap by 0.185 Jaccard ($0.028 \rightarrow 0.213$) while using $\sim 4.5 \times 10^{-5}$ kWh less per run—consistent with parts-based decompositions yielding sparse, interpretable structure on short, bag-of-words text [3]. LSI shows a lighter-weight version of the same story (+0.047 Cv, similar diversity, near-lowest energy), matching how truncated SVD exposes stable distributional patterns with minimal compute [10]. Among probabilistic and non-parametric models, HDP lifts coherence (+0.087 Cv) but increases redundancy (+0.290 Jaccard) and cuts diversity (0.240), reflecting known sensitivities of mixture models and coherence measures on sparse, short texts [2, 11, 12, 13]. On the embedding side, BERTopic most strongly improves separability (diversity +0.400 to 1.0; Jaccard 0.213 to 0.0) while maintaining solid coherence (+0.138 Cv), as expected from clustering sentence embeddings with class-TF-IDF refinement [4]; the trade-off is a clear energy premium ($+4.03 \times 10^{-4}$ kWh), echoing Green AI cautions [6, 5]. Top2Vec underperforms on coherence (0.171 Cv) and draws more energy ($+6.6 \times 10^{-5}$ kWh) [14]. Finally, GSDMM shows undefined coherence and the highest energy, underscoring a mismatch with intra-line code-switching [15].



LIMITATIONS



Corpus Representation

Drawn from publicly available Nigerian Lyrics

Likely Underrepresented niche sub-genres, older recordings, and artists without a digital footprint

Evaluation Metrics

Relies on intrinsic metrics instead of expert cultural judgements

Modelling Scope

Finite model and hyper parameter coverage

Other encoders or short-text methods might shift performance rankings

Linguistic & Structural Issues

Short, code-switched lines can blur topic boundaries

Line-level language IA: may mislabel or fragment coherences

Energy Measurements

Logged on a single hardware profile

Report in kWh only-no grid-based CO₂ estimates

Results may vary across and energy profiles

Copyright & Domain Boundaries

Only derived artifacts are released

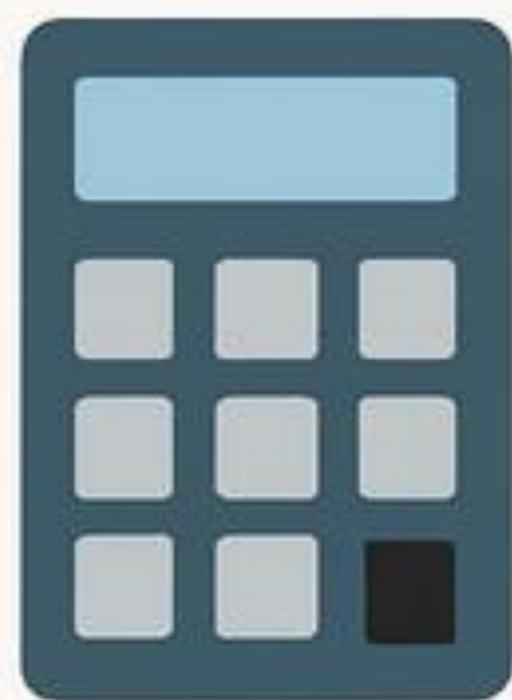
Findings apply primarily to short English-Yoruba-Pidgin lyrics: mother domains or languages may behave differently



Conclusion

Our results point to a clear, workable path for topic modelling on short, code-switched Nigerian lyrics.

Classical methods



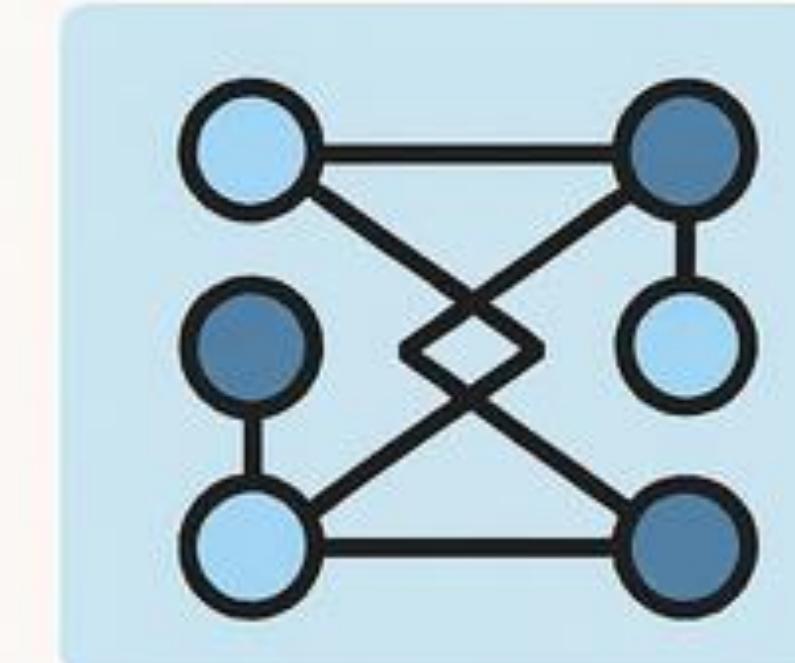
NMF

- High coherence
- Low overlap
- Tiny energy

While LSI provides a light baseline

This balance matters for classrooms, archives, and community projects running on limited power

Embedding pipelines



BERTopic delivers crisp separation
LDA remains frugal yet blurrier
Top2Vec underperforms



In practice

- Start with NMF (or LSI for speed)
- Monitor diversity and overlap
- Reserve BERTopic for strict disjointness



We thank the Artificial Intelligence Systems Research Group, Department of Computer Sciences, Olabisi Onabanjo University, and HERtificial Intelligence Research Group, Department of Mathematical and Computing Sciences, Thomas Adewumi University, Oko-Irese, Nigeria, and all contributors to Nigerian music archives.



Future work will combine classical and lightweight embedding methods with culturally grounded human evaluation.

Thank you for watching — and for advancing inclusive, sustainable AI.

