**TOWARD SENTIMENT ANALYSIS OF BURMESE MUSIC LYRICS: A TEXT-BASED ANALYSIS THROUGH MACHINE LEARNING AND DEEP LEARNING APPROACHES**

AYE KHIN KHIN HPONE

**I. INTRODUCTION**

Burmese song lyrics are a culturally significant form of artistic expression, encapsulating the emotions, stories, and traditions of Myanmar across different generations. Despite their cultural richness, there is limited computational research analyzing Burmese song lyrics for sentiment classification. While sentiment analysis has been widely applied to high-resource languages, Burmese remains underrepresented in natural language processing (NLP) research, with a scarcity of labeled datasets and linguistic resources.

This study focuses exclusively on **sentiment classification** of Burmese song lyrics, comparing the performance of traditional machine learning algorithms and advanced deep learning techniques. By building a structured dataset and evaluating model performance across multiple tasks, the study aims to contribute to the growing field of **Burmese computational linguistics** and create foundational work that could support future research in **digital humanities**, **language modeling**, and **cultural preservation**.

**II. OBJECTIVES**

* To develop a clean and structured dataset of at least **1,000 Burmese song (lyrics)**
* To apply **sentiment analysis** to the lyrics
* To compare the performance of Machine Learning (e.g. **Naive Bayes)** and **deep learning** on the Burmese text classification tasks
* To contribute to the foundational development of **Burmese-language NLP tools and resources**, especially in the cultural and linguistic context

**III. SENTIMENT DEFINITION**

Sentiment in this study refers to the predominant emotional polarity conveyed in the entire lyric, as interpreted by the annotator.

Positive: Expressions of joy, love, hope, pride, inspiration, and optimism.

Negative: Expressions of sadness, heartbreak, anger, regret, disappointment, or despair.

Neutral: Lyrics with no strong emotional tone, factual descriptions, or a balance of positive and negative emotions.

**IV. DATASET**

**Population & Sample**

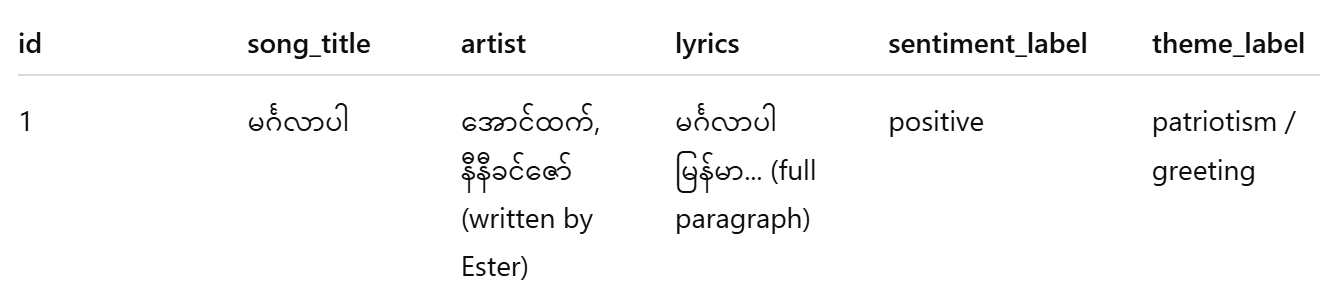
* **Population**: All Burmese song lyrics from traditional, modern, and contemporary genres.
* **Sample**: A minimum of **1,000 lyrics**, ensuring diversity in era, genre, and emotional tone.

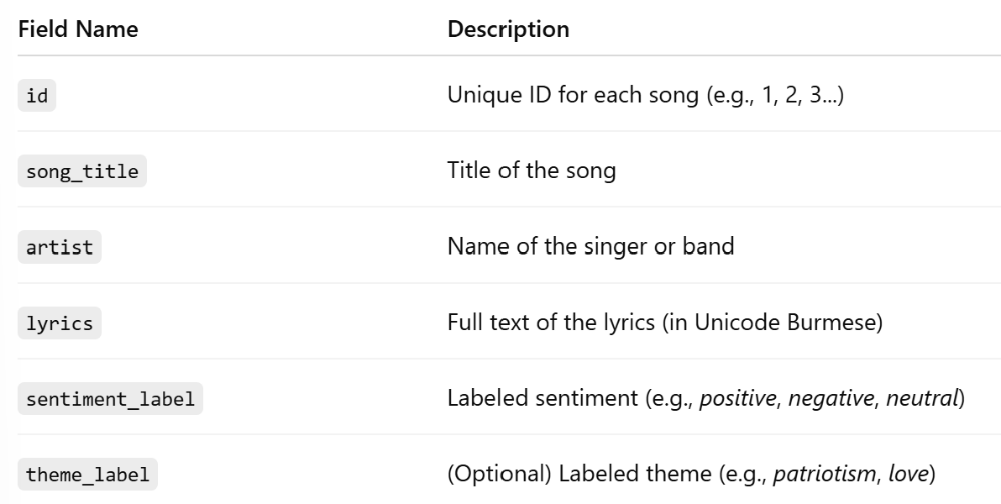
**Data Sources**

1. **Websites**:

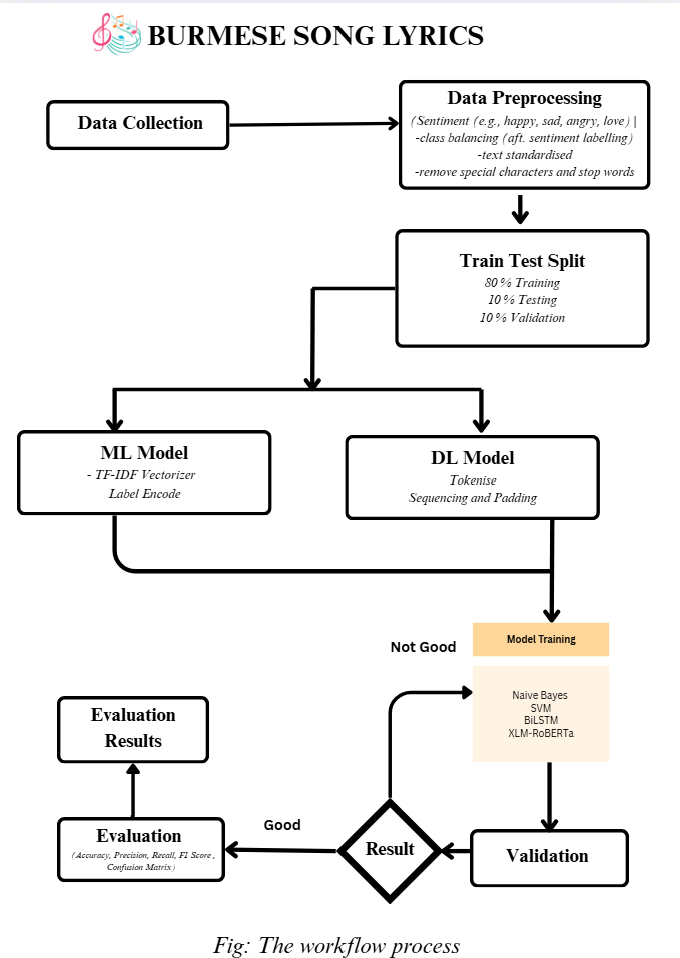
* [**LyricsTranslate**](https://lyricstranslate.com/en/songs) – An international lyrics database that includes Burmese songs with English translations, helpful for semantic validation.
* [**MyanmarLyrics.net**](https://myanmarlyrics.net/) – A dedicated Burmese music lyrics repository offering a large collection across traditional and modern genres.
* [**Facebook Page: Myanmar Song Lyrics and Chords**](https://www.facebook.com/MyanmarSongLyricsandChords/) – A community-driven source for both popular and lesser-known songs, often shared in textual format or graphic images.
* **Photographs of printed lyrics** – Lyrics sourced from album booklets, posters, or physical music collections will also be transcribed for inclusion in the dataset.

1. **Printed/Scanned Sources**:
   * Album booklets, music magazines, posters.
   * **OCR extraction** for Burmese Unicode text using Google’s OCR API and other open-source OCR tools.





**V. METHODOLOGY**



**VI. EXPERIMENT**

To demonstrate how the proposed system processes and classifies Burmese lyrics, we present a paragraph-style lyric excerpt from the song **"မင်္ဂလာပါ (Mingalaba)"**, performed by ​**အောင်ထက်(Aung Htet)** and နီနီခင်ဇော် **(Ni Ni Khin Zaw)** and written by **Ester**. This song exemplifies themes of national pride, beauty, and positive identity, making it ideal for both sentiment and theme classification.

မင်္ဂလာပါ နှုတ်ဆက်ခြွေလို့ သာယာဖွယ် မနက်ခင်းလေး အစပြုတယ်...

ရွှေဗဟိုရ်စည်သံ မိုးထက်ညံလို့ ပြည်ရွာမကျန် အင်အားအသစ်တို့ရခဲ့

လှပတဲ့ တောတောင်တွေနဲ့ စိမ်းလန်းတဲ့ မွေးဌာနီရယ်

ဘုရားပုထိုးနဲ့ ကျောင်းကန်စုံညီ မြင့်မြတ်တဲ့ သမိုင်းရာဇဝင်နဲ့ အခိုင်အမာရပ်တည်ထားတယ်

မြန်မာ မြန်မာ မြန်မာ မြန်မာ

မြန်မာလို နှုတ်ဆက်ပါရစေ မင်္ဂလာပါ...

|  |  |
| --- | --- |
| Expected Output | |
| Classification Task | Output |
| Sentiment | Positive |
| Theme | Patriotism / Greeting / Identity |

VII. EXPECTED OUTCOME

