ARTIST AND GENRE CLASSIFICATION FROM LYRICS

Vedika Chauhan, Tania Guerrero, Josue Martinez

Introduction

Music is a universal language, deeply linked to the origins of culture and communication. However, the massive volume and continuous evolution of music can make genre classification a difficult task. We address this problem by proposing a classification system that predicts the genre of the artist of a song. We trained both a Multitask classifier with a multilayer neural network and LSTM (Long Short-Term Memory) models using the same datasets to predict genre and artist. We also propose song-generation models based on a given artist and key keywords as input.

Methods

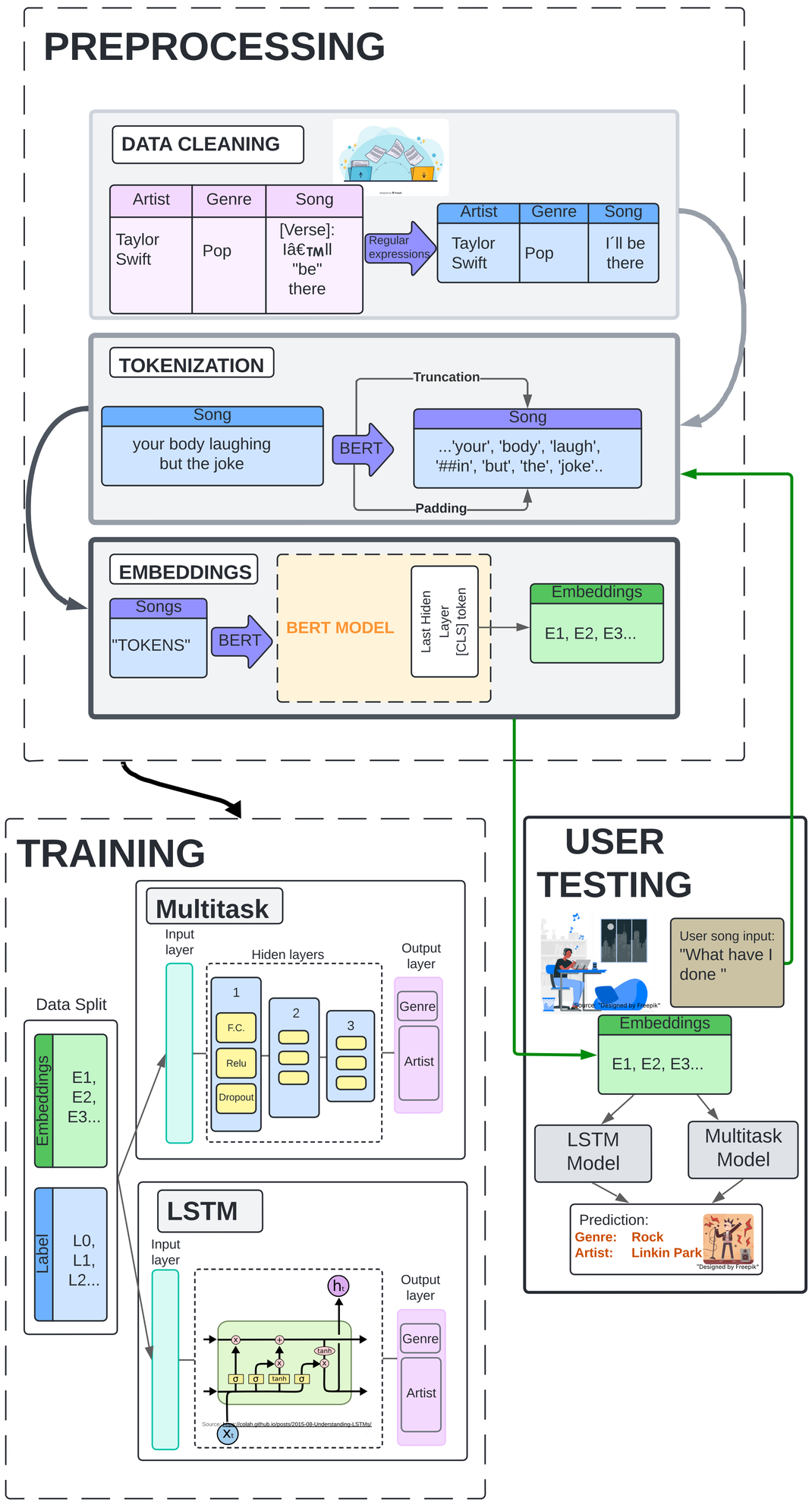

Figure 1: Classification process of music genres and artists

The accuracy of the prediction model is directly related to the size of the dataset.

Pre-processing

The implementation uses a dataset from Kaggle [1], which contains a collection of approximately 5 million songs and poems in various languages up to the year 2022. Figure 2 illustrates the artist and the genre used to train the models. The process was divided into three phases pre-processing, model design, and evaluation.

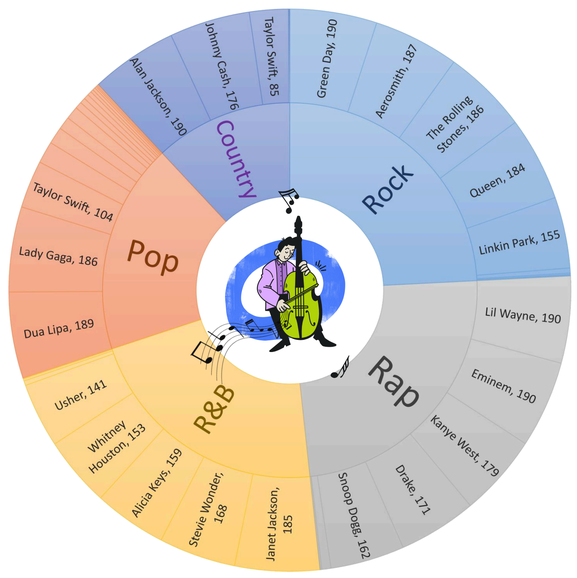

Figure 2: Dataset Overview: Artists by Genre

During the pre-processing phase, only English songs were selected from the data set. Special characters and punctuation were then removed, and Unicode characters were replaced by their ASCII equivalent.

Classification

We proposed two models for song classification: a Long Short-Term Memory (LSTM) based model, which is defined with a single LSTM layer followed by two fully connected layers for genre and artist classification. The second is a Multitask model, which has three fully connected layers with ReLU activation and dropout for regularization. Both models were trained over 350 epochs using the Adam optimizer with a learning rate of 0.0005.

Song Generator

The Song Generator is a tool that creates new song lyrics in the style of Taylor Swift. It works by analyzing Taylor Swift's top songs and breaking down the lyrics into smaller parts. The process begins with data preprocessing, which involves loading and filtering lyrics, followed by text cleaning and creating a corpus. A FastText model is trained to generate word embeddings, which are then used to prepare sequences for training. These sequences are loaded into a DataLoader to train in mini-batches instead of using the entire dataset in one training pass. The core of the model is an LSTM-based neural network, which is defined and trained with specific hyperparameters, loss functions, and optimizers. Finally, the model generates lyrics using temperature sampling and formats them into verses and choruses to produce lyrical output.

Experimentation and Results

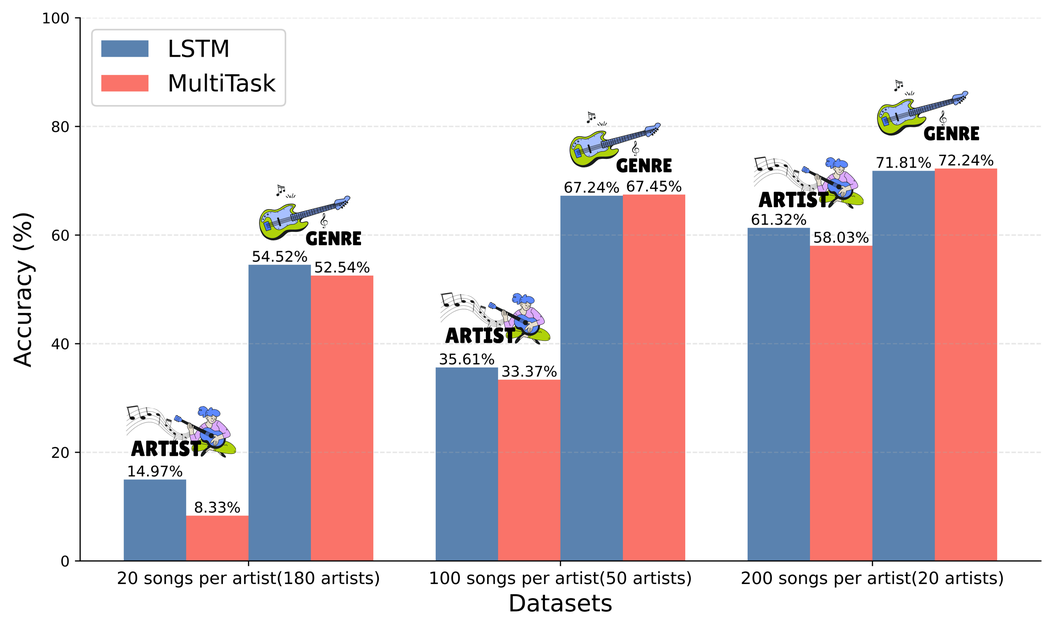

Figure 3: Accuracy Assessment of LSTM and MultiTask Models

The models were trained with approx 800 songs of each, Rap, Rock, Pop, Country and R&B. The evaluation process involved calculating the accuracy of the test set by comparing the predicted labels with the actual labels. In Figure 3 can be observed that both models shows a general trend of increasing accuracy with larger datasets, for genre and artist predictions.

References

[1] Carlos Garcia and Kaggle Users. Genius song lyrics with language information. https://www.kaggle.com/datasets/carlosgdcj/genius-song-lyrics-with-language-information/data, 2024. [Accessed 24-June-2024].

[2] Tania Guerrero Vedika Chauhan, Josue Martinez. Nlp project. https://github.com/TaniaGithub0401/NLP_project.git, 2024. Accessed: 2024-06.

Vedika Chauhan
Hochschule Bonn-Rhein-Sieg
Email: vchauh2s@smail.inf.h-brs.de

Josue Martinez
Hochschule Bonn-Rhein-Sieg
Email: dmarti2s@smail.inf.h-brs.de

Tania Guerrero
Hochschule Bonn-Rhein-Sieg
Email: tguerr2s@smail.inf.h-brs.de

b-it Bonn-Aachen International Center for Information Technology

Hochschule Bonn-Rhein-Sieg
University of Applied Sciences