

# Ensemble RNN Model

## Model Overview

The Ensemble RNN model consists of four sub-models, each designed to process a different type of input data: Log Mel Spectrogram, Spectrogram, Chromagram, and Waveform. These sub-models are combined to make predictions about the genre of music based on multi-input data. The ensemble model aims to leverage the strengths of each individual sub-model to improve overall classification accuracy.

## Sub-models

1. **Log Mel Spectrogram Model:** Processes log mel spectrogram data using a recurrent neural network (RNN) architecture with Gated Recurrent Units (GRU).
2. **Spectrogram Model:** Processes spectrogram data using a recurrent neural network (RNN) architecture with Long Short-Term Memory (LSTM) units.
3. **Chromagram Model:** Processes chromagram data using a recurrent neural network (RNN) architecture with GRU units.
4. **Waveform Model:** Processes waveform data using a recurrent neural network (RNN) architecture with LSTM units.

## Input Data

The input data for the ensemble model consists of four types:

- **Log Mel Spectrogram:** A representation of the audio signal's frequency content over time.
- **Spectrogram:** A visual representation of the audio signal's frequency content over time.
- **Chromagram:** A representation of the audio signal's chroma features.
- **Waveform:** The raw audio waveform data.

## Output

The model predicts the genre of music for each input sample. The output is one of four classes: Classical, Pop, R&B, or Rock.

## Performance

The performance of the ensemble model is evaluated based on its accuracy in predicting the correct genre label for each input sample. Performance metrics such as accuracy are calculated for individual sub-models as well as the ensemble model as a whole.

## Usage

The ensemble model can be used to make predictions on new music samples by providing the corresponding log mel spectrogram, spectrogram, chromagram, and waveform data as input. The model utilizes selective activation to determine which sub-models to use for each input sample based on their performance on training data.

## Limitations and Ethical Considerations

- The model's performance may vary depending on the quality and diversity of the training data.
- It's important to consider potential biases in the training data and their impact on model predictions, particularly in the context of music genre classification.
- The model may not generalize well to music genres or styles that are underrepresented in the training data.
- Care should be taken when interpreting the model's predictions, and human judgment should be used to verify the accuracy of the classifications.

## Maintenance

The model may require periodic updates or retraining to maintain its performance as new data becomes available or changes in the distribution of music genres occur. Regular monitoring of model performance and evaluation against ground truth labels is recommended to ensure continued reliability.

## Disclaimer

This model card is provided for informational purposes only. The authors make no guarantees regarding the accuracy or reliability of the model's predictions, and users should exercise caution and critical judgment when interpreting its output.