Review of: "Songs Classification Problem Research by Genre Based on Neural Network"

Rizwan Patan[1]

1 Kennesaw State University

Potential competing interests: No potential competing interests to declare.

- While MFCC is a valuable feature, explore other audio features like chroma, spectral centroid, and spectral roll-off. These can capture different aspects of music and potentially improve classification accuracy.
- Consider creating new features by combining existing ones or applying transformations. For example, combining MFCC with chroma might capture both timbral and harmonic information more effectively.
- Given the temporal nature of audio signals, CNNs can be particularly effective. Experiment with different CNN architectures, such as VGG, ResNet, or DenseNet, to capture patterns in the audio data.
- For sequential data like audio, RNNs can be beneficial. Explore variants like Long Short-Term Memory (LSTM) or Gated Recurrent Units (GRU) to capture long-term dependencies.
- To increase the dataset size and improve generalization, apply these techniques to create variations of the audio samples.
- Simulate real-world noise conditions by adding background noise to the audio. This can help the model become more robust to variations in recording quality.
- Combine different audio samples to create new, synthetic samples. This can introduce diversity and help the model learn more complex patterns.