BOULDERMUSICML : MUSIC GENRE CLASSIFICATION

TEA WRIGHT, ISABELLE MEREDITH, & SAMRIDHI SHRESTHA

SUMMARY

OUR MACHINE LEARNING PROJECT AIMS TO SOLVE THE CHALLENGE OF GENRE CLASSIFICATION FOR SONGS, WHICH IS DIFFICULT DUE TO THE DIVERSE AND OVERLAPPING CHARACTERISTICS OF DIFFERENT GENRES. THE PROPOSED SOLUTION INVOLVES CREATING AN ALGORITHM THAT CAN RECOGNIZE PATTERNS AND CLASSIFY SONGS BASED ON THEIR DEFINING FEATURES. THE UNIQUE ASPECT OF THIS APPROACH IS THE USE OF REAL BACKGROUND NOISE TO COMPARE THE ACCURACY OF THE MODEL BETWEEN CLEAN AND NOISY DATA.

OUR APPROACH

- COMBINED A CONVOLUTIONAL NEURAL NETWORK (CNN) AND K-NEAREST NEIGHBORS (KNN) CLASSIFICATION ALGORITHM.
- WE TESTED THEIR APPROACH BY SPLITTING THE DATA INTO TRAINING, VALIDATION, AND TEST DATA USING SKLEARN'S 'TRAIN_TEST_SPLIT' FEATURE.
- AS A RESULT, WE GOT AN ACCURACY SCORE OF 0.29 WITHOUT ANY FEATURE SELECTION AND 0.65 AFTER FEATURE SELECTION.
- IMPLEMENTED LIBROSA TO OVERLAY BACKGROUND NOISE AND COMPARED THEIR RESULTS TO THE BASELINE ACCURACIES.

DATA

- THE GTZAN DATASET PROVIDES ONE THOUSAND SONGS, EACH WITH DESCRIPTIVE DATA, GENRE, AND A SPECTROGRAM.
- THE DATASET HAS TEN GENRES WITH 100 SONGS PER GENRE.
- TO REDUCE BIAS IN THE DATASET, LOG SCALING WAS USED IN PREPROCESSING THE SPECTROGRAMS USING THE LIBROSA LIBRARY.
- FOUR TYPES OF BACKGROUND NOISE WERE OVERLAID ON THE LOG-SCALED SPECTROGRAMS FOR COMPARISON PURPOSES.
- THIS RESULTED IN 1000 LOG-SCALED SPECTROGRAMS AND 4000 OVERLAID PLOTS FOR EACH TYPE OF BACKGROUND NOISE.

RESULTS

- USED KNN AND CNN CLASSIFIERS ON CLEAN SOUND RECORDING DATASET
- BASELINE ACCURACY OF 0.65 ACHIEVED WITH KNN AFTER FEATURE SCALING AND SELECTION
- CNN ACHIEVED 0.6183 ACCURACY WITH PLAIN SPECTROGRAM DATA AFTER 500 EPOCHS
- LOG SCALING SPECTROGRAMS IMPROVED ACCURACY TO 0.6195
- TESTED MODEL'S ROBUSTNESS WITH VARYING LEVELS OF BACKGROUND NOISE
- ACCURACY DECREASED FOR MOST INSTANCES OF OVERLAID NOISE COMPARED TO CLEAN DATA
- "SPRING" BACKGROUND NOISE PRODUCED HIGHER ACCURACY OF 0.6558, HIGHLIGHTING IMPORTANCE OF TESTING ON VARIOUS TYPES OF DATA.

Mel-frequency spectrogram

Mel-frequency spectrogram