

Speech Commands: A Dataset for Limited-Vocabulary Speech Recognition

Report Summary

The Speech Commands dataset contains 105,829 utterances from 2,618 speakers, with audio files in WAVE format. It's divided into training, validation, and testing sets, ensuring consistent evaluation and avoiding overfitting. Evaluation metrics include matched percentage, wrong percentage, and false positives. The dataset enables comparisons across models and has been widely used in research on speech recognition, noise tolerance, and adversarial attacks.