

Dataset Name: YouTube Movie Trailer Dataset

Description: This dataset contains features of video and audio content obtained from movie trailers on YouTube. Each example is from a movie trailer.

Data Source: Videos obtained as a result of searches made with certain keywords via YouTube API.

Attributes:

- rate: The trailer's IMDB score. **(output)**
- image: Histograms of visual features obtained from the video content of the trailer.

RGB Histograms: Separate histograms for red, green and blue channels. The average of each(R-G-B) was recorded.
- sound: Features obtained from the audio content of the trailer.

Statistical Features from Audio Spectrogram:

- ♦ std_pow: Standard deviation of power levels in the spectrogram.
- ♦ max_pow: Maximum power level.
- ♦ min_pow: Minimum power level.
- ♦ mean_pow: Average of power levels.
- ♦ max_power_freq: Frequency of the highest power level.
- ♦ max_power_time: Time of the highest power level.
- ♦ Band Features (band5, band10, ..., band_last): Average power levels for specific frequency bands in the spectrogram.

Data Format: Files created separately for each trailer as CSV files. ';' There are columns separated by characters.

Example Usage Areas:

- To analyze the visual and audio content of movie trailers.
- Data preparation for movie recommendation systems.
- Making predictions from trailer content with machine learning or deep learning models.

Notes:

- The dataset is preprocessed and features are extracted.
- The data set is in an labeled (supervised) structure.

Prepared by:

- 201805010-Hüseyin Tüğyan
- 201805014-Beyza Kartal
- 201805046-Gamze Nur Çulha
- 211805004-İlke Doğan
- 211805078-Ali Cemal Gülmez
- 211805118-Engin Halil Yedirmeç