بسم الله الرحمن الرحیم

شرکت مهندسی نرم افزاری هلو

**گزارش شناسایی پخش موسیقی در یک فایل صوتی**

کاری از امیرعلی نسیمی

**شبکه عصبی SampleCNN**

مدل SampleCNN یک شبکه عصبی کانولوشنی (Convolutional Neural Network یا CNN) است که برای طبقه‌بندی فایل‌های صوتی طراحی شده است. این مدل از چندین لایه کانولوشنی تشکیل شده است که در ادامه توضیح داده می‌شوند:

**جریان کار پردازش صوت**

* main():
* تعیین مسیر فایل صوتی `test.wav`.
* ایجاد نمونه‌ای از کلاس `Music\_Detection`.
* اندازه‌گیری زمان شروع و پایان تحلیل فایل صوتی.
* چاپ زمان اجرا و درصد استفاده از حافظه.

M**usic\_Detection.py**

1. \_load\_model():

* بارگذاری پارامترهای مدل از فایل `model-gztan-speech-music-20000.pth`.
* تنظیم مدل به حالت ارزیابی (evaluation).

1. analyze():

* تبدیل فایل صوتی به فرمت 16 کیلوهرتز و مونو با استفاده از ffmpeg.
* بارگذاری فایل تبدیل شده به صورت بلوک‌های زمانی با استفاده از librosa.
* پردازش هر بلوک با مدل `SampleCNN` و چاپ نتیجه (موسیقی یا گفتار) به همراه احتمال.

**پردازش صوت**

1. **تبدیل فایل صوتی:**

* فایل ورودی با استفاده از ffmpeg به فرمت 16kHz و مونو تبدیل می‌شود.

1. **بارگذاری بلوک‌های صوتی:**

* فایل تبدیل شده به بلوک‌های زمانی با طول 59049 نمونه و گام 16000 نمونه تقسیم می‌شود.
* هر بلوک به عنوان یک ورودی برای مدل استفاده می‌شود.

1. **پردازش بلوک‌های صوتی:**

* هر بلوک به مدل `SampleCNN` وارد می‌شود.
* مدل احتمال اینکه بلوک حاوی موسیقی یا گفتار باشد را محاسبه می‌کند.
* نتایج همراه با زمان هر بلوک چاپ می‌شود.

**خلاصه جریان کار**

1. اسکریپت `run.py` اجرا می‌شود.
2. فایل صوتی `test.wav` بارگذاری و مسیر آن تعیین می‌شود.
3. نمونه‌ای از کلاس `Music\_Detection` ایجاد می‌شود.
4. فایل صوتی به فرمت مناسب تبدیل می‌شود.
5. فایل تبدیل شده به بلوک‌های زمانی تقسیم می‌شود.
6. هر بلوک توسط مدل `SampleCNN` پردازش می‌شود.
7. نتایج پردازش (موسیقی یا گفتار) همراه با احتمال مربوطه چاپ می‌شود.
8. زمان اجرای تحلیل و درصد استفاده از حافظه چاپ می‌شود.

**نتیجه‌گیری**

این سیستم یک مدل یادگیری عمیق برای تشخیص موسیقی و گفتار از فایل‌های صوتی است که از چندین لایه کانولوشنی برای استخراج ویژگی‌ها و یک لایه کاملاً متصل برای طبقه‌بندی استفاده می‌کند. جریان کار شامل تبدیل فایل صوتی به فرمت مناسب، تقسیم آن به بلوک‌های زمانی و پردازش هر بلوک با مدل `SampleCNN` است.