

مقدمه

طبقه‌بندی موسیقی بر اساس ژانر و سبک همواره یکی از موضوعاتی بوده که در حوزه هوش مصنوعی و یادگیری ماشین بسیار مورد توجه قرار گرفته است. شناسایی سبک یک موسیقی می‌تواند درک ما را از محتوای صوتی تا حد قابل توجهی افزایش دهد و اطلاعات ارزشمندی را در اختیار ما قرار دهد. یکی از رویکردهایی که معمولاً در زمینه دسته‌بندی‌های صوتی (Sound Classification) مورد استفاده قرار می‌گیرد، استفاده از Hidden Markov Models (HMM) می‌باشد. HMM یک مدل آماری است که یک چارچوب برای مدل‌سازی داده‌های متوالی ارائه می‌دهد و به عنوان یک روش مناسب برای تحلیل پدیده‌های متغیر بر اساس زمان به مانند موسیقی شناخته می‌شود. هدف اصلی نیز در این پروژه توسعه یک سیستم دسته‌بندی سبک موسیقی بر اساس HMM می‌باشد. برای اینکار مجموعه‌ای صوتی که دارای لیبل سبک موسیقی می‌باشند در اختیارتان قرار گرفته و شما باید مراحل پیش‌پردازش، استخراج ویژگی، پیاده‌سازی مدل و تحلیل آن را در طی پروژه انجام دهید.

آشنایی با مجموعه داده

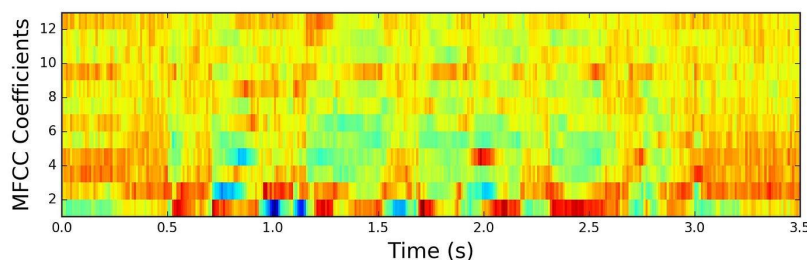
مجموعه داده‌ای که در این پروژه در اختیارتان قرار گرفته، مجموعه‌ای شامل ۴ سبک موسیقی متفاوت می‌باشد که هر کدام دارای ۱۰۰ فایل صوتی هستند. همچنین تمام فایل‌های صوتی داده شده با فرمت wav در اختیارتان قرار گرفته‌اند و زمان تقریبی ۳۰ ثانیه‌ای دارند. چهار سبکی که در این پروژه از آن‌ها استفاده شده است، شامل pop, metal, hip-hop و blues می‌باشد. مجموعه داده را می‌توانید از [این لینک](#) دریافت کنید.

پیش پردازش و استخراج ویژگی

داده‌هایی که در اختیارتان قرار گرفته‌اند، تا حد خوبی پیش‌پردازش شده‌اند. برای مثال ابتدا و انتهای فایل‌های موسیقی معمولاً مقداری خالی می‌باشد که در اینجا چنین قسمت‌هایی حذف شده‌اند. همچنین همان‌طور که گفته شد، زمان تمام فایل‌ها به صورت تقریبی ۳۰ ثانیه می‌باشد اما دقیقاً یکسان نمی‌باشد. یکی دیگر از روش‌هایی که در پردازش داده‌های صوتی استفاده می‌شود، شکستن هر صوت به قسمت‌های کوچک‌تر با اندازه یکسان می‌باشند. با استفاده از چنین روشی نه تنها زمان تمام نمونه‌ها یکسان می‌شود، بلکه تعداد نمونه‌ها نیز افزایش خواهد یافت که می‌تواند در شرایطی که تعداد نمونه داده‌ها محدود می‌باشد بسیار کمک‌کننده باشد و همچنین دقت مدل ما را افزایش دهد.

۱. به چه علت نیاز می‌باشد که نمونه‌ها دارای طول یکسان باشند؟

استخراج و انتخاب ویژگی از مهم‌ترین مراحل هر پروژه هوش مصنوعی و یادگیری ماشین می‌باشند. استفاده از ویژگی مناسب می‌تواند تاثیر بالایی در خروجی مدل نهایی داشته باشد. همچنین ویژگی‌های بسیار متفاوتی را می‌توان از یک محتوای صوتی استخراج کرد که هر کدام اطلاعات گوناگونی را به ما می‌دهند. در این پروژه از ضرایب MFCC (Mel Frequency Cepstral Coefficient) برای دسته‌بندی موسیقی‌ها استفاده می‌کنیم. ضرایب MFCC به طور گسترده در زمینه‌های مرتبط با دسته‌بندی موسیقی و تشخیص گفتار استفاده شده‌اند.



در این قسمت مجموعه ضرایب MFCC را برای هر کدام از نمونه‌های داده شده استخراج نمایید. برای این کار می‌توانید از یکی از کتابخانه‌های موجود برای کار با صوت استفاده نمایید. پس از استخراج ویژگی‌ها، نمودار Heat Map مربوط به ضرایب برای یک نمونه از هرکدام از دسته‌ها رسم نمایید.

2. چرا در محاسبه MFCC فریم‌های استفاده شده با یکدیگر هم‌پوشانی دارند؟
3. چرا در اکثر پروژه‌های مرتبط با صوت تنها از ۱۲ یا ۱۳ ضریب ابتدایی MFCC استفاده می‌شود؟

آشنایی با HMM

همانطور که گفته شد، HMM یک مدل آماری است که به عنوان یک روش مناسب برای تحلیل پدیده‌های متغیر بر اساس زمان شناخته می‌شود. هر مدل HMM از بخش‌های زیر تشکیل شده است:

States

Observations

Transition Probabilities

Emission Probabilities

1. توضیح دهید منظور از State ها و Observation چیست؟ در این تمرین State ها کدامند و Observation چگونه بدست می‌آید؟
2. مدل‌های HMM را میتوان بر اساس میزان وابستگی میان State های پنهان دسته‌بندی کرد، مدلی که در این تمرین به پیاده‌سازی آن می‌پردازید یک مدل First-Order HMM است. دلیل نامگذاری آن و همچنین ویژگی‌های آن را بررسی کنید و تفاوت آن با مدل‌های دیگر در این دسته‌بندی را بیان کنید.
3. درباره HMM تحقیق کنید و توضیح دهید که این مدل برای بررسی و تحلیل چه پدیده‌هایی مناسب است؟ چرایی این موضوع را توضیح دهید.

4. مدل HMM نیز مانند هر مدل دیگری دارای مزایا و معایبی است که آن را ویژه می‌کند. مزایا و معایب این مدل را بررسی کرده و هر کدام را مختصراً توضیح دهید.
5. انواع مختلفی از مدل های HMM وجود دارد، درباره آنها تحقیق کنید و چند مورد را بطور مختصر بررسی کنید.

پیاده‌سازی مسئله

بخش اول: Implementing with Libraries

در بخش اول، شما باید با استفاده از کتابخانه‌های آماده (hmmlearn) به طراحی و پیاده‌سازی یک مدل HMM بپردازید. سپس مدل را بر روی داده‌هایی که در اختیار دارید Train کرده و در نهایت نتایج را با معیارهایی که در بخش ارزیابی و تحلیل معرفی می‌شوند، بررسی کنید.

بخش دوم: Implementing from Scratch

در این بخش از تمرین به پیاده‌سازی یک مدل HMM از پایه می‌پردازید. یک فایل Notebook شامل کدهای ناقص مورد نیاز برای پیاده‌سازی این مدل برای شما آپلود شده است. با تکمیل بخش‌های مختلف این فایل، در نهایت به یک مدل HMM دست خواهید یافت. مدل پایانی را بر روی داده‌ها train کرده و نتایج آن را با معیارهای متفاوتی که در ادامه معرفی می‌شوند، بررسی و تحلیل کنید.

راهنمایی پیاده‌سازی:

_state_likelihood:

این متد احتمال observation-ها را برای هر state محاسبه می‌کند. می‌توانید برای تکمیل این بخش از Multivariate normal distribution استفاده کنید.

_em_step:

ابتدای این متد مقادیر جدید alpha و beta محاسبه می‌شوند تا در الگوریتم Expectation-Maximization به کار گرفته شوند.

Expectation-Maximization شامل آپدیت کردن پارامترهای مدل از جمله:

B (observation probability),

transition_matrix (transition probability),

initial_prob (states probability at the first time)

می‌باشد. همچنین میانگین و کوواریانس هم دوباره محاسبه میشوند که این دو مورد به طور کامل پیاده سازی شده‌اند و نیاز به تکمیل ندارند.

train:

مدل ما در فرایند train در چرخه re-estimate و Expectation-Maximization قرار می‌گیرد تا likelihood دیدن observation-های آن مدل بیشتر شود.

score:

با استفاده از این متد احتمال دیده شدن یک observation را در مدل به دست می‌آوریم.

ارزیابی و تحلیل

معیارهای زیادی برای سنجش و ارزیابی عملکرد مدل‌ها وجود دارد. در این تمرین از شما می‌خواهیم برای ارزیابی مدل خود از ۴ معیار زیر استفاده کنید:

Accuracy

Precision

Recall

F1 Score

درباره هر کدام از معیارهای بالا تحقیق کنید و نحوه محاسبه هر یک را توضیح دهید. مطالعه این دو لینک ([لینک ۱](#) و [لینک ۲](#)) برای درک این موضوع به شما کمک خواهد کرد.

1. توضیح دهید که هر کدام از معیارها چگونه مدل را ارزیابی میکنند.
2. تفاوت میان Precision و Recall را بیان کنید و توضیح دهید چرا هر کدام به تنهایی برای ارزیابی مدل کافی نیست؟ برای هر یک مثالی بیاورید که در آن، این معیار مقدار بالایی دارد اما مدل عملکرد خوبی ندارد.
3. معیار F1 از چه نوع میانگین‌گیری استفاده میکند؟ تفاوت این نوع میانگین‌گیری با میانگین‌گیری عادی چیست و در اینجا چرا اهمیت دارد؟
4. همانطور که می‌دانید، در این مسئله بیش از ۲ کلاس داریم، در مورد Multi-Class Metrics تحقیق کنید و سه حالت میانگین‌گیری Macro و Micro و Weighted را شرح دهید. برای مطالعه می‌توانید از این [لینک](#) استفاده کنید.
5. برای نتایج بدست آمده، Confusion Matrix رسم کنید و معیارهای بالا را برای هر کلاس به صورت جداگانه و سپس با استفاده از سه نوع میانگین‌گیری گفته شده برای تمام کلاس‌ها محاسبه کنید.
6. مقادیر بدست آمده برای معیارهای ارزیابی را تحلیل کنید.
7. ممکن است نتایج شما در بخش اول و دوم فرق کند و مدل آماده (که در بخش اول از آن استفاده کردید) نتایج متفاوت و دقت بالاتری نسبت به مدل طراحی شده توسط شما داشته باشد. این اختلاف ممکن است چه دلایلی داشته باشد؟ درباره عوامل تاثیرگذار بر روی این اختلاف دقت تحقیق کنید. راهنمایی: این تفاوت می‌تواند در ساختار مدل‌ها یا پیش‌پردازش داده‌ها باشد.

نکات پایانی

- توضیحات مربوط به هر بخش از پروژه را بطور خلاصه و در عین حال مفید در گزارش خود ذکر کنید. از ابزارهای تحلیل داده مانند نمودارها استفاده کنید.
- پس از مطالعه کامل و دقیق صورت پروژه، در صورت وجود هرگونه ابهام یا سوال با طراحان پروژه در ارتباط باشید.
- نتایج، گزارش و کدهای خود را در قالب یک فایل فشرده با فرمت AI_CA2_[stdNumber].zip در سامانه ایلرن بارگذاری کنید.
- محتویات پوشه باید شامل فایل jupyter-notebook، خروجی html و فایل‌های مورد نیاز برای اجرای آن باشد. از نمایش درست خروجی‌های مورد نیاز در فایل html مطمئن شوید.
- دقت کنید که نیازی به آپلود مجموعه داده‌ها در سامانه ایلرن نیست.