

حسام رمضانیان

گزارش پروژه سوم هوش مصنوعی

پیش پردازش و استخراج ویژگی

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

علت یکسان بودن سایز نمونه‌ها این است که باید از یک فرمت استاندارد پیروی کنند که طول یکسان هم یکی از این شاخص‌ها است باید استاندارد شود زیرا داده‌های موجود در نمونه طولانی تر باعث اختلال در فرایند تحلیل داده‌ها می‌شود چون همتایی برای آن‌ها در نمونه کوتاه تر وجود ندارد

(۲) چرا در محاسبه MFCC فریم‌های استفاده شده با یکدیگر همپوشانی دارند؟

همپوشانی فریم‌ها باعث می‌شود که انتقال بین ضرایب mfcc به صورت پیوسته باشد و همچنین از دست رفتن داده‌ها در هنگام انتقال از یک فریم به فریم دیگر کاهش یابد و داده‌های mfcc بهتری را از یک فایل در اختیار خواهیم داشت

(۳) چرا در اکثر پروژه‌های مرتبط با صوت تنها از ۱۲ یا ۱۳ ضریب ابتدایی MFCC استفاده میشود؟

دلیل استفاده از ۱۲ یا ۱۳ ضریب اول این است که ضرایب پایین تر داده‌های مهم‌تری را در اختیار ما قرار می‌دهند و ضرایب بالاتر داده‌هایی مانند نویزها را در اختیار ما قرار می‌دهند همچنین با توجه به اینکه ۱۲ یا ۱۳ ضریب اول تقریباً تمامی اطلاعات مورد نیاز ما برای محاسبات را فراهم می‌کند تنها از همین ضرایب استفاده می‌کنیم تا زمان و توان محاسباتی کمتری را نیاز داشته باشیم

آشنایی با HMM

(۱) توضیح دهید منظور از State ها و Observation چیست؟ در این تمرین State ها کدامند و Observation چگونه بدست می‌آید؟

Observation استیت‌هایی هستند که ما امکان مشاهده آنها را داریم و با استفاده از آنها هیدن استیت موجود را استنباط می‌کنیم

در این تمرین observation ها همان ضرایب استخراج شده از هر فایل هستند

state یا همان هیدن استیت به حالت‌هایی گفته می‌شود که ما امکان مشاهده آنها را نداریم می‌خواهیم در یک مدل **hmm** آنها را با استفاده از **Observation** ها استنباط کنیم

در این پروژه هیدن استیت‌ها مفهوم خاصی ندارند و صرفاً یک سری استیت فرضی هستند که با تغییر آنها به تعداد استیت مورد نیاز برای حل مسئله می‌رسیم

(۲) 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 **HMM Order-First** است. دلیل نامگذاری آن و همچنین ویژگی‌های آن را بررسی کنید و تفاوت آن با مدل‌های دیگر در این دسته‌بندی را بیان کنید.

دلیل این نامگذاری این است که احتمال یک استیت در زمان تنها به زمان قبلی وابسته است و ویژگی که در بالا ذکر شد باعث می‌شود که احتمالات را برای هر حالت بتوان به صورتی محاسبه کرد که تنها به حالت قبلی وابسته است در واقع میتوانیم با توان محاسباتی کمتر مدل را پیاده‌سازی کنیم

حالات دیگر **hmm** مانند مرتبه صفر مرتبه ۲ یا مرتبه‌های بالاتر وجود دارد در حالت مرتبه صفر ترنژیشن بین استیت‌ها وجود ندارد و در واقع احتمال هر استیت به هیچ استیت دیگری وابسته نیست همچنین در مرتبه‌های بیش از ۱ احتمال هر استیت به بیش از یک استیت قبل وابسته است

(۳) درباره **HMM** تحقیق کنید و توضیح دهید که این مدل برای بررسی و تحلیل چه پدیده‌هایی مناسب است؟ چرایی این موضوع را توضیح دهید .

در انواع مسایلی که قرار است از مشاهده یک پدیده در طول زمان پدیده دیگر که قابل مشاهده نیست اما وابستگی بین این دو وجود دارد را پیدا کنیم می‌توان از **hmm** کمک گرفت برای مثال پردازش صوت مدل سازی‌های مالی و

(۴) 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

- معایب **hmm** ها برای استنتاج استیت های پنهان تنها از مشاهدات استفاده می‌کند و اثر متغیرهای خارجی را نادیده می‌گیرند همچنین در صورتی که اندازه مدل بزرگ باشد تخمین پارامترهای آنها نیز مشکل است

- مزایا : این مدل ها می توانند با تنها با استفاده از استیت های مشاهده شده استیت های پنهان را شناسایی کند ، همچنین می توانند با داده های بسیار کم الگوهای پیچیده ای را محاسبه کنند و می توان با تغییراتی آن را به سادگی با تعداد زیادی از مسائل سازگار کرد

۵) انواع مختلفی از مدل های HMM وجود دارد، درباره آنها تحقیق کنید و چند مورد را بطور مختصر بررسی کنید

HMM انواع متنوعی دارند که عبارتند از: گاوسی، پیوسته، گسسته، چندبعدی و سلسله مراتبی.

گاوسی: این مدل یکی از انواع خاص HMM است که در آن هر دو متغیر پنهان و مشاهده شده به صورت متغیرهای تصادفی پیوسته در نظر گرفته می شوند. از آنجایی که بسیاری از سیگنال های واقعی مانند سیگنال صوتی یا سری های زمانی مالی، ذاتاً پیوسته هستند، استفاده از مدل گاوسی می تواند نمایش بهتر و دقیق تری از این سیگنال ها فراهم کند.

پیوسته: متغیرهای پنهان به صورت یک مجموعه پیوسته در نظر گرفته می شوند. این امر انعطاف پذیری بالاتری را فراهم می آورد. هنگامی که استیت های سیستم ذاتاً پیوسته باشند، استفاده از این مدل مناسب است.

گسسته: مدل گسسته متداول ترین نوع است که در آن هم متغیرهای پنهان و هم مشاهده شده به صورت گسسته در نظر گرفته می شوند.

تحلیل و ارزیابی :

۱) توضیح دهید که هر کدام از معیارها چگونه مدل را ارزیابی میکنند.

accuracy: فرمول آن به صورت $(TP+TN)/ALL$ معیار مناسبی برای زمانی است که کلاس های ما تعداد متفاوتی نمونه دارند در غیر این صورت معمولاً معیار مناسبی نیست

Recall: در سوال ۲ توضیح داده شده

Precision: در سوال ۲ توضیح داده شده است

F1-score = $2 \times (\text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$ فرمول آن به این صورت است

در این معیار ضعف های دو معیار قبل بر طرف شده و فقط زمانی مقدار آن زیاد است که هر دو معیار قبل مقدار زیادی داشته باشند

۲) تفاوت میان Recall و Precision 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

recall: فرمول آن به این صورت است $(TP/(TP+FN))$ در این معیار تنها با افزایش FN مقدار کاهش میابد در صورتی که FP ها در نظر گرفته نمیشوند

مثال: یک مدل تشخیص بیماری از بین ۱۰۰ نفر که ۵ نفر آن ها بیمار هستند ۵ نفر را بیمار اعلام می کند و ۹۵ نفر را غیر بیمار اما recall این وضعیت بسیار کم است

Precision: فرمول این معیار به این صورت است $(TP/(TP+FP))$ در این معیار با افزایش FP مقدار کاهش میابد در صورتی که FN ها در نظر گرفته نمیشوند

مثال: دوباره مثال بالا را در نظر میگیریم با این تفاوت که از ۹۵ فرد سالم ۵ نفر بیمار تشخیص داده شدند در این صورت مقدار این معیار ۵۰ درصد می شود در صورتی که ۹۰ نفر به درستی سالم تشخیص داده شده اند

۳) معیار F_1 از چه نوع میانگینگیری استفاده میکند؟ تفاوت این نوع میانگینگیری با میانگینگیری عادی چیست و در اینجا چرا اهمیت دارد؟

F_1 از میانگین هارمونیک استفاده میکند زیرا میانگین هارمونیک باعث می شود که مقدار بالاتری را برای حالت هایی در نظر بگیریم که هر دو معیار recall و precision مقدار مناسبی دارند و برای حالت هایی که فقط یکی از آن ها وضعیت خوبی دارد مقدار کمتری را داشته باشیم اما میانگین عادی در صورت زیاد بودن فقط یک معیار هم مقدار بالایی را نشان می دهد

۴) همانطور که میدانید، در این مسئله بیش از ۲ کلاس داریم، در مورد Metrics Class-Multi تحقیق کنید و سه حالت میانگینگیری Macro و Micro و Weighted را شرح دهید

Macro-average: این متریک میانگین عملکرد را به ازای هر کلاس محاسبه می کند. سپس، میانگین این مقادیر را به عنوان عملکرد کلی گزارش می دهد.

Weighted-average: این متریک مانند micro-average عمل می کند، با این تفاوت که به هر کلاس بر اساس فراوانی آن کلاس در داده ها وزن می دهد.

Micro-average: این متریک در واقع برای همه معیارها ثابت است و مقدار TPها را به کل نمونه به دست می آورد

(۵) برای نتایج بدست آمده، **Matrix Confusion** رسم کنید و معیارهای بال را برای هر کالس به صورت جداگانه و سپس با استفاده از سه نوع میانگین گیری گفته شده برای تمام کالس ها محاسبه کنید.

در فایل report.html موجود است

(۶) مقادیر بدست آمده برای معیارهای ارزیابی را تحلیل کنید.

accuracy: همانطور که مشاهده میشود این معیار معیار مناسبی برای سنجش نیست زیرا مقدار آن برای ژانر blues در مدلی که خودمان پیاده سازی کردیم مقدار نسبتاً بالایی است در صورتی که مدل هیچ یک از تست های این ژانر را به درستی پیشبینی نمی کند

Recall: این معیار برای سنجش بهتر از معیار قبلی عمل میکند زیرا برای blues مقدار صفر را نشان می دهد همچنین برتری هر مدل در ژانر ها مختلف را به درستی مشخص می کند

Precision: این معیار نیز برای سنجش مناسب نیست زیرا همانطور که مشاهده می شود برتر مدل ها در تشخیص ژانر را در نظر نمی گیرد برای مثال در ژانر hiphop با وجود اینکه مدلی که خودمان پیاده سازی کردیم عملکرد بهتری در تشخیص دارد اما مقدار آن در این معیار کمتر است اما این معیار ضعف معیار قبل که تشخیص ندادن تعداد پیشبینی های نادرست از یک ژانر هست را ندارد

F1: این معیار بهترین معیار برای سنجش است زیرا به درستی برتری مدل ها را مشخص می کند و ضعف هر دو معیار قبل را برطرف میکند و با مشاهده آن میتوان برتری مدل کتابخانه را متوجه شد

(۷) 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 اختالف ممکن است چه دالیلی داشته باشد؟ درباره عوامل تاثیرگذار بر روی این اختالف دقت تحقیق کنید. راهنمایی: این تفاوت میتواند در ساختار مدلها یا پیشپردازش دادهها باشد.

با توجه به اینکه پیش پردازش داده ها یکسان است ممکن است علت اختالف در الگوریتم مورد استفاده در هر یک از مدل ها باشد که سبب ایجاد نتایج بهتر یا بدتر می شود.